

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:13:56 ; Search time 137 Seconds  
(without alignments)  
5197.647 Million cell updates/sec

Title: US-09-576-989-3

Perfect score: 10465

Sequence: 1 MAPITASQTRGLGCIIT.....PMWCLLLSVGVGIIILPNR 1985

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Sep04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	10465	100.0	1985	5	AAO18001	Aao18001 Hepatitis
2	10465	100.0	1985	5	AAE15717	Aae15717 Hepatitis
3	10465	100.0	1985	8	ADJ57846	Adj57846 HCV repli
4	10462	100.0	1985	5	AAE15729	Aae15729 Hepatitis
5	10462	100.0	2201	5	ABG30581	Abg30581 Hepatitis
6	10462	100.0	2201	5	ABG30587	Abg30587 Hepatitis
7	10462	100.0	3010	5	ABG32451	Abg32451 Hepatitis
8	10460	100.0	1985	5	AAE15731	Aae15731 Hepatitis
9	10460	100.0	1985	5	AAE15730	Aae15730 Hepatitis
10	10459	99.9	1985	5	AAE15727	Aae15727 Hepatitis
11	10458	99.9	1985	5	AAE15728	Aae15728 Hepatitis
12	10458	99.9	3010	5	ABG32460	Abg32460 Hepatitis
13	10458	99.9	3010	5	ABG32454	Abg32454 Hepatitis
14	10457	99.9	3010	5	ABG32461	Abg32461 Hepatitis
15	10456	99.9	3010	5	ABG32458	Abg32458 Hepatitis
16	10456	99.9	3010	5	ABG32459	Abg32459 Hepatitis
17	10456	99.9	3010	5	ABG32455	Abg32455 Hepatitis
18	10456	99.9	3010	5	ABG32457	Abg32457 Hepatitis
19	10456	99.9	3010	5	ABG32452	Abg32452 Hepatitis
20	10455	99.9	1985	5	AAE15720	Aae15720 Hepatitis
21	10455	99.9	3010	5	ABG32453	Abg32453 Hepatitis
22	10451.5	99.9	3011	5	ABG32456	Abg32456 Hepatitis
23	10450	99.9	1985	5	AAE15722	Aae15722 Hepatitis
24	10447	99.8	2201	5	ABG30582	Abg30582 Hepatitis
25	10442	99.8	2201	5	ABG30580	Abg30580 Hepatitis

26	10441	99.8	2201	5	ABG30594	Abg30594 Hepatitis
27	10439	99.8	2201	5	ABG30584	Abg30584 Hepatitis
28	10439	99.8	2201	5	ABG30589	Abg30589 Hepatitis
29	10438	99.7	2201	5	ABG30591	Abg30591 Hepatitis
30	10438	99.7	2201	5	ABG30593	Abg30593 Hepatitis
31	10438	99.7	2201	5	ABG30590	Abg30590 Hepatitis
32	10437	99.7	2201	5	ABG30598	Abg30598 Hepatitis
33	10437	99.7	2201	5	ABG30595	Abg30595 Hepatitis
34	10436	99.7	2201	5	ABG30601	Abg30601 Hepatitis
35	10435	99.7	2201	5	ABG30602	Abg30602 Hepatitis
36	10435	99.7	2201	5	ABG30599	Abg30599 Hepatitis
37	10434	99.7	2201	5	ABG30585	Abg30585 Hepatitis
38	10434	99.7	2201	5	ABG30600	Abg30600 Hepatitis
39	10428	99.6	2201	5	ABG30583	Abg30583 Hepatitis
40	10421	99.6	2201	5	ABG30586	Abg30586 Hepatitis
41	10421	99.6	2201	5	ABG30588	Abg30588 Hepatitis
42	10392	99.3	2063	7	ADD67963	Add67963 Hepatitis
43	10202	97.5	1985	6	ABU09575	Abu09575 HCV Met-N
44	10181	97.3	1985	6	ABU09574	Abu09574 HCV Met-N
45	10170	97.2	2201	2	AAW01680	AAW01680 HCV NS2-N

## ALIGNMENTS

### RESULT 1

AAO18001  
XX AAO18001 standard; protein; 1985 AA.  
AC AAO18001;  
XX  
DT 30-AUG-2002 (first entry)  
XX  
DE Hepatitis C virus NS3 proteinase/helicase.  
XX  
KW Hepatitis; HCV; core-neo; NS3 proteinase/helicase; vaccine; diagnosis;  
KW virucide; hepatotropic; gene therapy; anti-viral; enzyme.  
XX  
OS Hepatitis C virus.  
XX  
PN WO200238793-A2.  
XX  
PD 16-MAY-2002.  
XX  
PF 02-NOV-2001; 2001WO-US046350.  
XX  
PR 07-NOV-2000; 2000US-0245866P.  
XX  
PA (ANAD-) ANADYS PHARM INC.  
XX  
PI Bichko V;  
XX  
DR WPI; 2002-490082/52.  
XX N-PSDB; AAL47276.  
XX  
FT Novel nucleic acid encoding replication competent recombinant hepatitis C  
virus genome useful for screening anti-hepatitis C virus therapeutics and  
for vaccine development.  
XX  
PS Claim 6; Page 50-51; 85pp; English.

The present invention provides protein and coding sequences from Hepatitis C virus (HCV), comprising all or part of the HCV genome and able to replicate efficiently when transfected into a susceptible cell line without reducing the growth rate of the cell line by more than 10 fold. The sequences are useful for screening for anti-HCV therapeutics, for detecting antibodies to HCV in a biological sample such as blood, serum, plasma, blood cells, lymphocytes, or liver cells from a subject, for deriving authentic HCV components such as replication-competent non-infectious, replication-defective infection-component, and replication-defective non-infectious HCV, in gene therapy or gene vaccination targeted to hepatic tissue for treating an animal infected or susceptible to HCV infection and for studying HCV infection and propagation. The

CC present sequence is the HCV NS3 proteinase/helicase

XX Sequence 1985 AA;

Query Match 100.0%; Score 10465; DB 5; Length 1985;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	M	A	P	I	T	A	S	Q	O	T	R	L	L	G	C	I	T	S	L	T	G	R	N	Q	V	E	G	V	Q	V	S	T	A	T	Q	S	F	L	A	T	C	V	N	G	V	C	W	T	V	H	G	A	60					
DB	1	M	A	P	I	T	A	S	Q	O	T	R	L	L	G	C	I	T	S	L	T	G	R	N	Q	V	E	G	V	Q	V	S	T	A	T	Q	S	F	L	A	T	C	V	N	G	V	C	W	T	V	H	G	A	60					
QY	61	G	S	K	T	L	A	G	P	K	P	I	T	Q	M	T	N	V	D	D	L	V	G	M	A	P	P	G	A	S	L	T	P	C	T	G	S	S	D	L	V	T	R	H	A	D	V	I	P	V	R	R	120						
DB	61	G	S	K	T	L	A	G	P	K	P	I	T	Q	M	T	N	V	D	D	L	V	G	M	A	P	P	G	A	S	L	T	P	C	T	G	S	S	D	L	V	T	R	H	A	D	V	I	P	V	R	R	120						
QY	121	G	D	S	R	G	S	L	L	P	R	P	V	S	Y	L	K	G	S	S	G	P	L	C	P	S	G	H	A	V	G	I	F	R	A	A	V	C	T	R	G	V	A	K	A	V	D	P	V	P	S	E	M	T	M	180			
DB	121	G	D	S	R	G	S	L	L	P	R	P	V	S	Y	L	K	G	S	S	G	P	L	C	P	S	G	H	A	V	G	I	F	R	A	A	V	C	T	R	G	V	A	K	A	V	D	P	V	P	S	E	M	T	M	180			
QY	181	R	S	P	V	F	T	D	N	S	S	P	R	A	V	P	O	T	F	O	V	A	H	L	H	A	P	T	G	S	K	S	T	K	V	P	A	A	Y	A	O	G	K	V	L	V	N	P	S	V	A	T	L	G	F	240			
DB	181	R	S	P	V	F	T	D	N	S	S	P	R	A	V	P	O	T	F	O	V	A	H	L	H	A	P	T	G	S	K	S	T	K	V	P	A	A	Y	A	O	G	K	V	L	V	N	P	S	V	A	T	L	G	F	240			
QY	241	A	N	M	S	K	A	G	I	D	P	N	I	R	T	G	V	R	T	T	G	A	P	I	T	S	T	G	A	P	I	T	S	T	G	A	P	I	T	S	T	G	A	P	I	T	S	T	G	A	P	I	T	S	T	300			
DB	241	A	N	M	S	K	A	G	I	D	P	N	I	R	T	G	V	R	T	T	G	A	P	I	T	S	T	G	A	P	I	T	S	T	G	A	P	I	T	S	T	G	A	P	I	T	S	T	G	A	P	I	T	S	T	300			
QY	301	I	L	G	I	G	T	V	L	D	O	A	E	T	A	G	A	R	L	V	V	L	A	T	A	T	P	G	S	V	T	V	P	H	P	N	I	E	E	V	A	L	S	T	G	E	I	P	F	Y	K	A	I	P	E	T	360		
DB	301	I	L	G	I	G	T	V	L	D	O	A	E	T	A	G	A	R	L	V	V	L	A	T	A	T	P	G	S	V	T	V	P	H	P	N	I	E	E	V	A	L	S	T	G	E	I	P	F	Y	K	A	I	P	E	T	360		
QY	361	K	G	R	H	L	I	F	C	H	S	K	K	C	D	E	L	A	K	L	S	G	L	G	N	A	V	A	Y	R	G	L	D	V	S	I	P	T	S	G	D	V	I	V	A	T	A	L	M	T	G	F	420						
DB	361	K	G	R	H	L	I	F	C	H	S	K	K	C	D	E	L	A	K	L	S	G	L	G	N	A	V	A	Y	R	G	L	D	V	S	I	P	T	S	G	D	V	I	V	A	T	A	L	M	T	G	F	420						
QY	421	G	D	F	S	V	I	D	C	N	T	C	T	Q	T	V	T	F	S	L	D	T	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	480										
DB	421	G	D	F	S	V	I	D	C	N	T	C	T	Q	T	V	T	F	S	L	D	T	F	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	T	480											
QY	481	E	R	P	S	G	M	F	D	S	S	V	L	C	E	C	Y	D	A	G	A	W	E	L	T	P	A	E	T	S	V	R	L	A	V	L	N	T	P	G	L	P	V	C	D	H	L	E	F	W	E	S	V	T	G	L	540		
DB	481	E	R	P	S	G	M	F	D	S	S	V	L	C	E	C	Y	D	A	G	A	W	E	L	T	P	A	E	T	S	V	R	L	A	V	L	N	T	P	G	L	P	V	C	D	H	L	E	F	W	E	S	V	T	G	L	540		
QY	541	T	H	I	D	A	H	F	L	S	O	T	K	O	A	G	N	F	P	L	V	A	Y	O	A	T	C	A	R	A	G	A	P	P	S	W	O	M	K	C	L	I	R	L	K	P	T	L	H	G	P	T	L	600					
DB	541	T	H	I	D	A	H	F	L	S	O	T	K	O	A	G	N	F	P	L	V	A	Y	O	A	T	C	A	R	A	G	A	P	P	S	W	O	M	K	C	L	I	R	L	K	P	T	L	H	G	P	T	L	600					
QY	601	Y	R	L	G	A	V	Q	N	E	V	T	T	H	P	T	K	Y	I	M	A	C	M	S	A	D	L	E	V	V	T	S	T	V	L	V	G	V	L	A	A	L	A	A	Y	C	L	T	T	G	S	V	I	V	R	660			
DB	601	Y	R	L	G	A	V	Q	N	E	V	T	T	H	P	T	K	Y	I	M	A	C	M	S	A	D	L	E	V	V	T	S	T	V	L	V	G	V	L	A	A	L	A	A	Y	C	L	T	T	G	S	V	I	V	R	660			
QY	661	I	L	S	G	P	A	I	P	D	R	E	V	L	Y	R	E	P	D	E	M	E	C	A	S	H	I	P	Y	I	E	Q	G	M	L	A	E	O	P	K	O	K	A	I	G	L	L	O	T	A	T	K	O	A	E	A	720		
DB	661	I	L	S	G	P	A	I	P	D	R	E	V	L	Y	R	E	P	D	E	M	E	C	A	S	H	I	P	Y	I	E	Q	G	M	L	A	E	O	P	K	O	K	A	I	G	L	L	O	T	A	T	K	O	A	E	A	720		
QY	721	A	A	P	V	S	K	M	T	L	E	A	F	A	K	M	N	F	I	S	G	I	O	Y	L	A	G	I	S	T	P	G	N	P	A	I	A	S	L	M	A	F	T	A	S	I	T	S	P	L	T	O	H	780					
DB	721	A	A	P	V	S	K	M	T	L	E	A	F	A	K	M	N	F	I	S	G	I	O	Y	L	A	G	I	S	T	P	G	N	P	A	I	A	S	L	M	A	F	T	A	S	I	T	S	P	L	T	O	H	780					
QY	781	T	L	L	F	N	I	L	G	V	A	R	A	L	A	P	P	S	A	S	A	F	V	G	A	I	A	G	A	V	G	S	I	G	L	K	V	L	V	D	I	L	A	G	Y	G	A	G	A	L	V	A	840						
DB	781	T	L	L	F	N	I	L	G	V	A	R	A	L	A	P	P	S	A	S	A	F	V	G	A	I	A	G	A	V	G	S	I	G	L	K	V	L	V	D	I	L	A	G	Y	G	A	G	A	L	V	A	840						
QY	841	F	K	M	S	G	E	M	P	S	T	E	D	L	V	N	L	P	A	I	L	S	P	G	A	L	V	G	V	W	C	A	A	I	L	R	H	V	G	P	E	G	A	V	Q	M	N	L	I	A	P	A	S	R	900				
DB	841	F	K	M	S	G	E	M	P	S	T	E	D	L	V	N	L	P	A	I	L	S	P	G	A	L	V	G	V	W	C	A	A	I	L	R	H	V	G	P	E	G	A	V	Q	M	N	L	I	A	P	A	S	R	900				
QY	901	G	N	H	V	S	P	H	Y	P	E	S	D	A	A	R	V	T	Q	I	L	S	S	L	T	I	T	Q	L	K	R	L	H	O	M	I	N	E	D	C	S	T	P	C	S	G	S	M	L	R	D	V	M	D	W	I	C	960	
DB	901	G	N	H	V	S	P	H	Y	P	E	S	D	A	A	R	V	T	Q	I	L	S	S	L	T	I	T	Q	L	K	R	L	H	O	M	I	N	E	D	C	S	T	P	C	S	G	S	M	L	R	D	V	M	D	W	I	C	960	
QY	961	T	V	L	T	D	F	K	T	W	L	Q	S	K	L	P	R	L	P	G	V	P	F	F	S	C	O	R	G	Y	K	G	V	M	R	G	D	G	I	M	Q	T	T	C	P	C	G	A	I	T	H	V	K	N	G	S	M	R	1020



AAE15717  
 ID AAE15717 standard; protein; 1985 AA.  
 AC AAE15717;  
 DT 12-MAR-2002 (first entry)  
 DE Hepatitis C virus (HCV) replBartMan polyprotein.  
 KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
 KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
 KW replBartMan.  
 OS Hepatitis C virus.  
 XX WO200189364-A2.  
 XX 29-NOV-2001.  
 XX 23-MAY-2001; 2001WO-US016822.  
 XX 23-MAY-2000; 2000US-00576989.  
 XX (UNIW ) UNIV WASHINGTON.  
 XX Rice CM, Blight KJ;  
 XX WPI; 2002-066755/09.  
 XX N-PSDB; AAD25322.  
 XX  
 XX Hepatitis C virus variants having greater transfection efficiency and  
 PT ability to survive subpassage, useful as a vaccine for immunizing primate  
 PT to the virus, comprise non-naturally occurring viral sequences.  
 XX  
 XX Claim 14; Page 65; 174pp; English.  
 XX  
 XX The invention relates to Hepatitis C virus (HCV) variants which include  
 CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
 CC variants that have a transfection efficiency and ability to survive  
 CC subpassage greater than HCV that have wild-type polyprotein coding  
 CC regions. The polynucleotides of the invention are useful for identifying  
 CC a cell line that is permissive for infection with HCV and detecting  
 CC replication of HCV in cells of the cell line. They are also useful for  
 CC testing a compound for anti-viral properties and for inhibiting HCV  
 CC infection. They are also useful for the generation of defined HCV virus  
 CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
 CC attachment, penetration and entry, structure/function studies on HCV  
 CC proteins and RNA elements and identification of new antiviral targets, a  
 CC systematic survey of cell culture systems and conditions to identify  
 CC those that support wild-type and variant HCV RNA replication and particle  
 CC release, production of adaptive HCV variants capable of more efficient  
 CC replication in cell culture, production of HCV variants with altered  
 CC tissue or species tropism, establishment of alternative animal models for  
 CC inhibitor evaluation including those supporting HCV variant replication,  
 CC development of cell-free HCV replication assays, production of  
 CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
 CC derivatives as possible vaccine candidates, engineering of attenuated or  
 CC defective HCV derivatives for expression of heterologous gene products  
 CC for gene therapy and vaccine applications and for utilisation of the HCV  
 CC glycoproteins for targeted delivery of therapeutic agents to the liver  
 CC or other cell types with appropriate receptors. Vaccine comprising these  
 CC sequences is useful for inducing immunoprotection to HCV in a primate.  
 CC The present sequence is Hepatitis C virus (HCV) replBartMan polyprotein  
 XX  
 XX Sequence 1985 AA;  
 SQ  
 Query Match 100.0%; Score 10465; DB 5; Length 1985;  
 Best Local Similarity 100.0%; Pred No 0;  
 Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 1 MAPITAYSQTRGLGCIITSLTGRDRNQVEGVQVQVSTATQSLATCVNGVCWTYHGA 60  
 1 MAPITAYSQTRGLGCIITSLTGRDRNQVEGVQVQVSTATQSLATCVNGVCWTYHGA 60

QY	61	GSKTLAGPKG	ITQMYTNVDQDLVGMQAPPGARSLT	PCTCGSSDLXLVTRHADVPVRRR	120	
DB	61	GSKTLAGPKG	ITQMYTNVDQDLVGMQAPPGARSLT	PCTCGSSDLXLVTRHADVPVRRR	120	
QY	121	GDSRGLLSRP	SVYLGSSGRLCPGSHAVGIFRAAVCT	RGVAKAVDFVPVSMETTM	180	
DB	121	GDSRGLLSRP	SVYLGSSGRLCPGSHAVGIFRAAVCT	RGVAKAVDFVPVSMETTM	180	
QY	181	RSPVFTD	NSPPAVPQTFQVAHLHAPTGS	GSKSTKVPAAAYAAQGYKVLNLSVAATLGF	240	
DB	181	RSPVFTD	NSPPAVPQTFQVAHLHAPTGS	GSKSTKVPAAAYAAQGYKVLNLSVAATLGF	240	
QY	241	AYMSKAH	GIDPNIRTVRTITTGAPITYSTY	GVKFLADGCGSGGAYDIIICDECHSTDTT	300	
DB	241	AYMSKAH	GIDPNIRTVRTITTGAPITYSTY	GVKFLADGCGSGGAYDIIICDECHSTDTT	300	
QY	301	ILGIGTVL	DOAETAGARLVLATATPPGS	VTVPHPNTEEVALSSTGIBPFYKAIPIETI	360	
DB	301	ILGIGTVL	DOAETAGARLVLATATPPGS	VTVPHPNTEEVALSSTGIBPFYKAIPIETI	360	
QY	361	KGGRHLIF	CHSKKKCDELAALKSLGLN	VAYYRGDLVSVIPTSGDVIVVATDALMTGFT	420	
DB	361	KGGRHLIF	CHSKKKCDELAALKSLGLN	VAYYRGDLVSVIPTSGDVIVVATDALMTGFT	420	
QY	421	GDPSVID	CNTCVTQTVD	FSLDPTFTTITTTVPQDAVSR	QRRGRTGRGMIYRFVTPG	480
DB	421	GDPSVID	CNTCVTQTVD	FSLDPTFTTITTTVPQDAVSR	QRRGRTGRGMIYRFVTPG	480
QY	481	ERPSG	MPDSSVLCEDYDAGCANYELT	PAETSVRLRAYLNTPLPVCODHLEFWESVFTGL	540	
DB	481	ERPSG	MPDSSVLCEDYDAGCANYELT	PAETSVRLRAYLNTPLPVCODHLEFWESVFTGL	540	
QY	541	THIDA	FLPSOTKQAGDNFPYLVAYQ	ATVCARAQAPPSWDQMKCLRLKPTLHGSPPLL	600	
DB	541	THIDA	FLPSOTKQAGDNFPYLVAYQ	ATVCARAQAPPSWDQMKCLRLKPTLHGSPPLL	600	
QY	601	YRLGAV	QNEVTTTHPIKYIMACMS	ADLEVTTSTWLVGGVLAALAAAYCLTTSWIVGR	660	
DB	601	YRLGAV	QNEVTTTHPIKYIMACMS	ADLEVTTSTWLVGGVLAALAAAYCLTTSWIVGR	660	
QY	661	IILSG	EAIPDREVLVREFDEMEBCASH	LPYIEQGMQLAEQFKQKAIQLLOTAKQAEA	720	
DB	661	IILSG	EAIPDREVLVREFDEMEBCASH	LPYIEQGMQLAEQFKQKAIQLLOTAKQAEA	720	
QY	721	AAPV	ESKWRTLFAFWAKHWNFI	SGTQYLAGLSTLPGNPAIASLMFTA	SITSPLTQH	780
DB	721	AAPV	ESKWRTLFAFWAKHWNFI	SGTQYLAGLSTLPGNPAIASLMFTA	SITSPLTQH	780
QY	781	TLFLN	ILGGMVAQAQLAPPSAAS	AFVAGIAGAAVSGIGLKVLDILAGYGAGVAGALVA	840	
DB	781	TLFLN	ILGGMVAQAQLAPPSAAS	AFVAGIAGAAVSGIGLKVLDILAGYGAGVAGALVA	840	
QY	841	FKVMS	GEMPTEDLVNLLPAILSP	GLVGVVCAATILRRHVGE	GAVQWNNLIIFASR	900
DB	841	FKVMS	GEMPTEDLVNLLPAILSP	GLVGVVCAATILRRHVGE	GAVQWNNLIIFASR	900
QY	901	GNHVS	THVPVPSDAAARVTQIL	SSLTITQLKRLHQWINE	DCSTPCSSGSLRDVMDWIC	960
DB	901	GNHVS	THVPVPSDAAARVTQIL	SSLTITQLKRLHQWINE	DCSTPCSSGSLRDVMDWIC	960
QY	961	TVLTD	FKTLQSKLLPRLP	GVVFFSCQGYKGVWRGDG	IMQTTCPGCAQITGHVKNKSGMR	1020
DB	961	TVLTD	FKTLQSKLLPRLP	GVVFFSCQGYKGVWRGDG	IMQTTCPGCAQITGHVKNKSGMR	1020
QY	1021	IYGP	RTCSNTHGTPPINAYTT	GPCTPSPAPNYSRALMR	VAAEEYVEVTRVGDFFHVTGM	1080
DB	1021	IYGP	RTCSNTHGTPPINAYTT	GPCTPSPAPNYSRALMR	VAAEEYVEVTRVGDFFHVTGM	1080
QY	1081	TTDNV	KCPQVPAPEFTEVDG	VRHLRYAPACKELLRE	VTFLVGLNOYLVQSOLPCEPE	1140
DB	1081	TTDNV	KCPQVPAPEFTEVDG	VRHLRYAPACKELLRE	VTFLVGLNOYLVQSOLPCEPE	1140

1141 PDVAVLTSMITDPSHITAEAKRLARGSPPLSSASSQSAPSIKATCTTRHDSPAD 1200  
 1141 PDVAVLTSMITDPSHITAEAKRLARGSPPLSSASSQSAPSIKATCTTRHDSPAD 1200  
 1201 LLEANLLWRQEMGNITRVESKVVILDSFPLQAEEDEREVSVPALILRSRKFPPRAM 1260  
 1201 LLEANLLWRQEMGNITRVESKVVILDSFPLQAEEDEREVSVPALILRSRKFPPRAM 1260  
 1261 PIWARPDPNPPLESWKDPDYVPPVHGCPLPAPAKAPP:PPPRKRRTVVLSBSTVSSALA 1320  
 1261 PIWARPDPNPPLESWKDPDYVPPVHGCPLPAPAKAPP:PPPRKRRTVVLSBSTVSSALA 1320  
 1321 ELATKTFGSSSAVDSGTATASPOPSDDGAGSDVESYSPMPPLGEGPDPDSDGSW 1380  
 1321 ELATKTFGSSSAVDSGTATASPOPSDDGAGSDVESYSPMPPLGEGPDPDSDGSW 1380  
 1381 STVSEASDDVCCSSSYTWTCALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSAS 1440  
 1381 STVSEASDDVCCSSSYTWTCALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSAS 1440  
 1441 LRQKKVTFORLOVLDHYPDVILKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFYGA 1500  
 1441 LRQKKVTFORLOVLDHYPDVILKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFYGA 1500  
 1501 DVRLSSKAVNHRSVWKDLLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDLG 1560  
 1501 DVRLSSKAVNHRSVWKDLLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDLG 1560  
 1561 VRVCEKWLVDVYSTLPQAVMGSSYGFQYSPGQVFEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
 1561 VRVCEKWLVDVYSTLPQAVMGSSYGFQYSPGQVFEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
 1621 TVTENDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGQNGYRRCRAGSVLT 1680  
 1621 TVTENDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGQNGYRRCRAGSVLT 1680  
 1681 TSCGNTLTCLYKAAACRAKLODCTMLVCGDDLVCESAGTQDEASLRAPTEAMTRY 1740  
 1681 TSCGNTLTCLYKAAACRAKLODCTMLVCGDDLVCESAGTQDEASLRAPTEAMTRY 1740  
 1741 SAPPGPPPKPEYDLELITSCSSNVSVVAHDASGRVYVLTDRPTTPLARAWEATRHPTVN 1800  
 1741 SAPPGPPPKPEYDLELITSCSSNVSVVAHDASGRVYVLTDRPTTPLARAWEATRHPTVN 1800  
 1801 SWLGNTIMVAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPIQIOL 1860  
 1801 SWLGNTIMVAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPIQIOL 1860  
 1861 HGLSAPLSHSYSGEINRVASCLKLGVPPLRVHRARSVRARLLSQGGRAATCGKLYF 1920  
 1861 HGLSAPLSHSYSGEINRVASCLKLGVPPLRVHRARSVRARLLSQGGRAATCGKLYF 1920  
 1921 NNAVTKLKLTPIPAAQSOLDLSSWVFAGYSGGDIYHLSLRARPRWFPMWCLLLLSVGVIY 1980  
 1921 NNAVTKLKLTPIPAAQSOLDLSSWVFAGYSGGDIYHLSLRARPRWFPMWCLLLLSVGVIY 1980  
 1981 LLPNR 1985  
 1981 LLPNR 1985

RESULT 3  
 ADJ57846 standard; protein; 1985 AA.

XX ADJ57846;  
 XX AC  
 XX ADJ57846;  
 DT 06-MAY-2004 (first entry)  
 XX HCV replicon.  
 DE hepatitis C virus; HCV; Antiinflammatory; Hepatotropic; Virucide;  
 XX HCV replicon.  
 KW

XX Unidentified.  
 OS WO2004015131-A2.  
 XX 19-FEB-2004.  
 XX 12-AUG-2003; 2003WO-US025260.  
 XX 12-AUG-2002; 2002US-0402661P.  
 XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 XX Gao M, Lemm JA, O'Boyle DR, Nower P, Rigat K, Sun J;  
 PI WPI; 2004-180685/17.  
 XX N-PSDB; ADJ57845.  
 XX Use of hepatitis C virus assays or reporter assays, e.g. identifying a  
 PT compound that inhibits hepatitis C virus RNA replication or identifying a  
 PT compound that modulates the activity of a gene of interest.  
 XX Claim 4; SEQ ID NO 2; 45pp; English.  
 XX The present invention relates to the use of hepatitis C virus (HCV)  
 CC assays for identifying a compound that inhibits HCV RNA replication and  
 CC reporter assays for identifying a compound that modulates the activity of  
 CC a gene of interest. The assays are useful for identifying a compound that  
 CC inhibits HCV RNA replication or for identifying a compound that modulates  
 CC the activity of a gene of interest. The HCV assay is useful for high  
 CC throughput screening that quantifies both the amount of HCV RNA  
 CC replication inhibitory activity associated with a test compound and the  
 CC amount of cytotoxicity associated with the test compound. The compound is  
 CC useful for treating hepatitis C infection. Assays of the invention have  
 CC distinct advantages when compared to qRT-PCR or other methods in that  
 CC assays of the invention may take place in situ in a detergent based crude  
 CC cell lysate, which requires no further preparation prior to performing  
 CC the assays. The assays do not also involve numerous manipulations to add  
 CC or subtract reagents after addition of test compounds and are desirably  
 CC based on a viral protein which is required by the HCV replicon for  
 CC replication. The present sequence represents a HCV replicon used in the  
 CC assay of the invention.  
 XX Sequence 1985 AA;  
 QY Query Match 100.0%; Score 10465; DB 8; Length 1985;  
 Db Best Local Similarity 100.0%; Pred. No. 0;  
 QY Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 1 MAPITAYSOQTRGLLGCIIITSLTGRDRNQVEGEVQVVVSTATQSFATCVNGVCWTVYHGA 60  
 1 MAPITAYSOQTRGLLGCIIITSLTGRDRNQVEGEVQVVVSTATQSFATCVNGVCWTVYHGA 60  
 QY 61 GSKTLAGPKGPIITQMTNVNDQDLVGMQAPPGARSLTPTCTCGSSDLXLVTRHADVIPVRR 120  
 Db 61 GSKTLAGPKGPIITQMTNVNDQDLVGMQAPPGARSLTPTCTCGSSDLXLVTRHADVIPVRR 120  
 QY 121 GDSRGSLLSPRPVSYLKGSSGGLLCPSHAVGIFRAAVCTRGVAKAVDFPVESMETTM 180  
 Db 121 GDSRGSLLSPRPVSYLKGSSGGLLCPSHAVGIFRAAVCTRGVAKAVDFPVESMETTM 180  
 QY 181 RSPVFTDSSPPPAVPOTFOVAHLHAPTSGSKTKVPAAYAAQYKVLVLPNSVAATLFG 240  
 Db 181 RSPVFTDSSPPPAVPOTFOVAHLHAPTSGSKTKVPAAYAAQYKVLVLPNSVAATLFG 240  
 QY 241 AYMSKAHGDIPNIRTVGRTITTTGAPITYSTYGFADGGCGSGGAYDIIICDECHSTSTT 300  
 Db 241 AYMSKAHGDIPNIRTVGRTITTTGAPITYSTYGFADGGCGSGGAYDIIICDECHSTSTT 300  
 QY 301 ILGIGTVLDQAEATAGARLVVLTATATPGSVTVPHNIEEVALSGTGPFGKAIPIETI 360  
 Db 301 ILGIGTVLDQAEATAGARLVVLTATATPGSVTVPHNIEEVALSGTGPFGKAIPIETI 360

361 KGRHLIFCHSKKKDELAALSGLGNNAVAYRGLDVSVIPTSGDVIIVATDALMTGFT 420  
Db ||||| LROKKVTFDLQVLDDHYRDVLEKMKAKASTVKAKLLSVEEACKLTPHSAKSGYGA 1500  
361 KGRHLIFCHSKKKDELAALSGLGNNAVAYRGLDVSVIPTSGDVIIVATDALMTGFT 420  
Qy ||||| DVRLSSKAVNHISVWKLLLEDTEPTDITIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560  
421 GDFSVIDCNTCVTQVDFSLDPTFTTETTPQDAVSRRGRTRGRMGIRFVFTPG 480  
Db ||||| DVRLSSKAVNHISVWKLLLEDTEPTDITIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560  
421 GDFSVIDCNTCVTQVDFSLDPTFTTETTPQDAVSRRGRTRGRMGIRFVFTPG 480  
Qy ||||| VRCEKMAVDVSTLPAQVMGSSYGFQSPGQRFVFNWAKKCPMGFAYDTRCFDS 1620  
481 ERPSGMFSSVLCBDCYDAGCAWYELTPAETSVRLRAYLNTGLPVCQDHLFWFVSFTGL 540  
Db ||||| VRCEKMAVDVSTLPAQVMGSSYGFQSPGQRFVFNWAKKCPMGFAYDTRCFDS 1620  
481 ERPSGMFSSVLCBDCYDAGCAWYELTPAETSVRLRAYLNTGLPVCQDHLFWFVSFTGL 540  
Qy ||||| TVTENDIRVEESIVQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGYRRCRASGVLT 1680  
541 THIDAFLSQTKQAGDNFPYLVAQATVCARAQAPPPSDOMWCKLRLKPTLHGPTLL 600  
Db ||||| TVTENDIRVEESIVQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGYRRCRASGVLT 1680  
541 THIDAFLSQTKQAGDNFPYLVAQATVCARAQAPPPSDOMWCKLRLKPTLHGPTLL 600  
Qy ||||| YRLGAVQNEVTTTHPIIKYIMACMSADLEVVTSTWLVGGVLAALAYCLTTGSAWIVGR 660  
601 YRLGAVQNEVTTTHPIIKYIMACMSADLEVVTSTWLVGGVLAALAYCLTTGSAWIVGR 660  
Db ||||| YRLGAVQNEVTTTHPIIKYIMACMSADLEVVTSTWLVGGVLAALAYCLTTGSAWIVGR 660  
601 YRLGAVQNEVTTTHPIIKYIMACMSADLEVVTSTWLVGGVLAALAYCLTTGSAWIVGR 660  
Qy ||||| IILSGKPAIIPDREVLREFDEMEECASHLPYIEQGMOLABOFPKQKALIGLQATKQAE 720  
661 IILSGKPAIIPDREVLREFDEMEECASHLPYIEQGMOLABOFPKQKALIGLQATKQAE 720  
Db ||||| IILSGKPAIIPDREVLREFDEMEECASHLPYIEQGMOLABOFPKQKALIGLQATKQAE 720  
661 IILSGKPAIIPDREVLREFDEMEECASHLPYIEQGMOLABOFPKQKALIGLQATKQAE 720  
Qy ||||| AAPVVEKWTLEAFWAKHWNFTISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
721 AAPVVEKWTLEAFWAKHWNFTISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
Db ||||| AAPVVEKWTLEAFWAKHWNFTISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
721 AAPVVEKWTLEAFWAKHWNFTISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
Qy ||||| TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
781 TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
Db ||||| TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
781 TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
Qy ||||| FKVMGSEMPSTEDLVNLLPAILSPGALVGVCAAILRRHVGGEGAVQWNNRLIAPASR 900  
841 FKVMGSEMPSTEDLVNLLPAILSPGALVGVCAAILRRHVGGEGAVQWNNRLIAPASR 900  
Db ||||| FKVMGSEMPSTEDLVNLLPAILSPGALVGVCAAILRRHVGGEGAVQWNNRLIAPASR 900  
841 FKVMGSEMPSTEDLVNLLPAILSPGALVGVCAAILRRHVGGEGAVQWNNRLIAPASR 900  
Qy ||||| GNHVSPTHYPESDAAARVQIILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
901 GNHVSPTHYPESDAAARVQIILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
Db ||||| GNHVSPTHYPESDAAARVQIILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
901 GNHVSPTHYPESDAAARVQIILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
Qy ||||| TVLTDFTKWLQSKLLPRLPGVPFFSCQYKGVWVGDMQTTCPGCAQITGHVKNXSMR 1020  
961 TVLTDFTKWLQSKLLPRLPGVPFFSCQYKGVWVGDMQTTCPGCAQITGHVKNXSMR 1020  
Db ||||| TVLTDFTKWLQSKLLPRLPGVPFFSCQYKGVWVGDMQTTCPGCAQITGHVKNXSMR 1020  
961 TVLTDFTKWLQSKLLPRLPGVPFFSCQYKGVWVGDMQTTCPGCAQITGHVKNXSMR 1020  
Qy ||||| IVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALRWAAEEYVEVTRVGDHFVYTG 1080  
1021 IVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALRWAAEEYVEVTRVGDHFVYTG 1080  
Db ||||| IVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALRWAAEEYVEVTRVGDHFVYTG 1080  
1021 IVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALRWAAEEYVEVTRVGDHFVYTG 1080  
Qy ||||| TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKPFLREEVTVLVLNQYLVSQLPCEPE 1140  
1081 TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKPFLREEVTVLVLNQYLVSQLPCEPE 1140  
Db ||||| TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKPFLREEVTVLVLNQYLVSQLPCEPE 1140  
1081 TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKPFLREEVTVLVLNQYLVSQLPCEPE 1140  
Qy ||||| PDVAVLTSMLTDPHSHTAETAKERLARGPPSLASSASQLSAPSUKATCTTRHDSPAD 1200  
1141 PDVAVLTSMLTDPHSHTAETAKERLARGPPSLASSASQLSAPSUKATCTTRHDSPAD 1200  
Db ||||| PDVAVLTSMLTDPHSHTAETAKERLARGPPSLASSASQLSAPSUKATCTTRHDSPAD 1200  
1141 PDVAVLTSMLTDPHSHTAETAKERLARGPPSLASSASQLSAPSUKATCTTRHDSPAD 1200  
Qy ||||| LIEANLLMRQEMGNITRVESENKVLDSFPLQAEDEDEREVSVPALIRSRKFPFRAM 1260  
1201 LIEANLLMRQEMGNITRVESENKVLDSFPLQAEDEDEREVSVPALIRSRKFPFRAM 1260  
Db ||||| LIEANLLMRQEMGNITRVESENKVLDSFPLQAEDEDEREVSVPALIRSRKFPFRAM 1260  
1201 LIEANLLMRQEMGNITRVESENKVLDSFPLQAEDEDEREVSVPALIRSRKFPFRAM 1260  
Qy ||||| PIWARPDPYNPPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
1261 PIWARPDPYNPPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db ||||| PIWARPDPYNPPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
1261 PIWARPDPYNPPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Qy ||||| ELATKTFGSSSESAVDSGTATASPDQSDGAGSDVESYSMPPLGEPGDPDLSGWS 1380  
1321 ELATKTFGSSSESAVDSGTATASPDQSDGAGSDVESYSMPPLGEPGDPDLSGWS 1380  
Db ||||| ELATKTFGSSSESAVDSGTATASPDQSDGAGSDVESYSMPPLGEPGDPDLSGWS 1380  
1321 ELATKTFGSSSESAVDSGTATASPDQSDGAGSDVESYSMPPLGEPGDPDLSGWS 1380  
Qy ||||| STVSEASEDVVCCSMYSYTWTCALITPCAETKLPINALSNLRLHNLVATTSRGAS 1440  
1381 STVSEASEDVVCCSMYSYTWTCALITPCAETKLPINALSNLRLHNLVATTSRGAS 1440  
Db ||||| STVSEASEDVVCCSMYSYTWTCALITPCAETKLPINALSNLRLHNLVATTSRGAS 1440  
1381 STVSEASEDVVCCSMYSYTWTCALITPCAETKLPINALSNLRLHNLVATTSRGAS 1440  
Qy ||||| LROKKVTFDLQVLDDHYRDVLEKMKAKASTVKAKLLSVEEACKLTPHSAKSGYGA 1500

Db ||||| LROKKVTFDLQVLDDHYRDVLEKMKAKASTVKAKLLSVEEACKLTPHSAKSGYGA 1500  
Qy ||||| DVRLSSKAVNHISVWKLLLEDTEPTDITIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560  
Db ||||| DVRLSSKAVNHISVWKLLLEDTEPTDITIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560  
Qy ||||| VRCEKMAVDVSTLPAQVMGSSYGFQSPGQRFVFNWAKKCPMGFAYDTRCFDS 1620  
Db ||||| VRCEKMAVDVSTLPAQVMGSSYGFQSPGQRFVFNWAKKCPMGFAYDTRCFDS 1620  
Qy ||||| TVTENDIRVEESIVQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGYRRCRASGVLT 1680  
Db ||||| TVTENDIRVEESIVQCCDLAPARQAIRSLTERLYIGGPLTNSKGQNGYRRCRASGVLT 1680  
Qy ||||| TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLAVICESAGTOEDEASLFAFTAMTRY 1740  
Db ||||| TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLAVICESAGTOEDEASLFAFTAMTRY 1740  
Qy ||||| SAPPGDPPKPEYDLELITSCSSNVSVVAHDASGKRVVYLTDRPTTPLARAAWETARHTPVN 1800  
Db ||||| SAPPGDPPKPEYDLELITSCSSNVSVVAHDASGKRVVYLTDRPTTPLARAAWETARHTPVN 1800  
Qy ||||| SWLGNIIIMYAPTWARMLMTHFFSILLAOEQLKALDQCIYGACYSIEPLDLPQIIORL 1860  
Db ||||| SWLGNIIIMYAPTWARMLMTHFFSILLAOEQLKALDQCIYGACYSIEPLDLPQIIORL 1860  
Qy ||||| HGLSAFSLHSYSQGEINRVASCLRKLGVPPLRVWRHARSVRARLLSQGGRAATCGKYL 1920  
Db ||||| HGLSAFSLHSYSQGEINRVASCLRKLGVPPLRVWRHARSVRARLLSQGGRAATCGKYL 1920  
Qy ||||| NWAVRTKLTPIPAASQLDSSWVFVAGYSGGDIYHLSLRARPRWFWMLLLSVGVGIY 1980  
Db ||||| NWAVRTKLTPIPAASQLDSSWVFVAGYSGGDIYHLSLRARPRWFWMLLLSVGVGIY 1980  
Qy ||||| LIPNR 1985  
Db ||||| LIPNR 1985  
Db ||||| LIPNR 1985  
Db ||||| LIPNR 1985  
RESULT 4  
AAE15729  
ID AAE15729 standard; protein; 1985 AA.  
XX  
AC AAE15729;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Hepatitis C virus (HCV) replbBartMan polyprotein variant A1174S.  
XX  
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
KW replbBartMan; mutant; mutein; variant.  
XX  
OS Hepatitis C virus.  
OS Synthetic.  
XX  
FH Key Location/Qualifiers  
FT Misc-difference 1174 /note= "Wild type Ala substituted with Ser"  
ET  
XX WO200189364-A2.  
XX  
XX 29-NOV-2001.  
XX  
XX 23-MAY-2001; 2001WO-US016822.  
XX  
XX 23-MAY-2000; 2000JUS-00576989.  
XX (UNIW ) UNIV WASHINGTON.  
XX  
XX Rice CM, Blight KJ;

DR WPI; 2002-066755/09.  
XX Hepatitis C virus variants having greater transfection efficiency and  
PT ability to survive subpassage, useful as a vaccine for immunizing primate  
PT to the virus, comprise non-naturally occurring viral sequences.  
XX  
PS Claim 14; Page; 174pp; English.  
XX  
XX The invention relates to Hepatitis C virus (HCV) variants which include  
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
CC variants that have a transfection efficiency and ability to survive  
CC subpassage greater than HCV that have wild-type polypeptide coding  
CC regions. The polynucleotides of the invention are useful for identifying  
CC a cell line that is permissive for infection with HCV and detecting  
CC replication of HCV in cells of the cell line. They are also useful for  
CC testing a compound for anti-viral properties and for inhibiting HCV  
CC infection. They are also useful for the generation of defined HCV virus  
CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
CC attachment, penetration and entry, structure/function studies on HCV  
CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of heterologous gene products  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the liver  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
CC The present sequence is Hepatitis C virus (HCV) repliconMan polyprotein  
CC variant. Note: The present sequence is not shown in the specification but  
CC is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the  
CC specification  
XX  
SQ Sequence 1985 AA;  
Query Match 100.0%; Score 10462; DB 5; Length 1985;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLCIIITSLTGRDRNQVEGEVQVQVSTATQSFATCNGVCWTVYHGA 60  
DB 1 MAPITAYSQOTRGLLCIIITSLTGRDRNQVEGEVQVQVSTATQSFATCNGVCWTVYHGA 60  
QY 61 GSKTLAGPKGPTITQMTNVDDQVLVHQAPPGARSLTPTCTGSSDYLVTTHADVIPVRR 120  
DB 61 GSKTLAGPKGPTITQMTNVDDQVLVHQAPPGARSLTPTCTGSSDYLVTTHADVIPVRR 120  
QY 121 GDSRGLLPSPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM 180  
DB 121 GDSRGLLPSPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM 180  
QY 181 RSPVFTDNSPPAVPOTFQVAHLHAPTSGSKSTKVPAAYAAQYKVLVLPNSVAATLGF 240  
DB 181 RSPVFTDNSPPAVPOTFQVAHLHAPTSGSKSTKVPAAYAAQYKVLVLPNSVAATLGF 240  
QY 241 AYSKAGIDPNTTGTVRTTTCAPTSTYTGKFLADGGCGGAYDIIICDECHSDSTT 300  
DB 241 AYSKAGIDPNTTGTVRTTTCAPTSTYTGKFLADGGCGGAYDIIICDECHSDSTT 300  
QY 301 ILGIGTVLDAQETAGARLVLATATPGSVTVPHNIEEVALSGTGPYKAIPIETI 360  
DB 301 ILGIGTVLDAQETAGARLVLATATPGSVTVPHNIEEVALSGTGPYKAIPIETI 360  
QY 361 KGGRHLLIFCHSKKCDLAAKLSGLGLNNAVAYRGLDVSVIPTSGDVIVVATDALMTGFT 420  
DB 361 KGGRHLLIFCHSKKCDLAAKLSGLGLNNAVAYRGLDVSVIPTSGDVIVVATDALMTGFT 420

QY 421 GDFSDVIDCNTCVTQTVDFSLDPTFTTETTTVPQDAVSRSRORGRGTGRGRIYRFTVTPG 480  
DB 421 GDFSDVIDCNTCVTQTVDFSLDPTFTTETTTVPQDAVSRSRORGRGTGRGRIYRFTVTPG 480  
QY 481 ERPSGMFSSVLCBCYDAGCAWYELTTPAETSRLRAYLNTPLPVCODHLEFWSVFTGL 540  
DB 481 ERPSGMFSSVLCBCYDAGCAWYELTTPAETSRLRAYLNTPLPVCODHLEFWSVFTGL 540  
QY 541 THIDAHFLSQTQAGDNPPYLVAQATVCARAQAPPPSDQMWKCLIRLKLKTLHGPTLL 600  
DB 541 THIDAHFLSQTQAGDNPPYLVAQATVCARAQAPPPSDQMWKCLIRLKLKTLHGPTLL 600  
QY 601 YRLGAVQNEVTTTHPIIKYIMACMSADLEVTSTWLVGGVLAALAAVCLITGSGVIVGR 660  
DB 601 YRLGAVQNEVTTTHPIIKYIMACMSADLEVTSTWLVGGVLAALAAVCLITGSGVIVGR 660  
QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMOLAEQFKOKAIGLQATKQAEA 720  
DB 661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMOLAEQFKOKAIGLQATKQAEA 720  
QY 721 AAPVVEKWRITLPAFWAKHMNFISGQIYLAGLSTLPGNPAIASLMAFTASITSLTTOH 780  
DB 721 AAPVVEKWRITLPAFWAKHMNFISGQIYLAGLSTLPGNPAIASLMAFTASITSLTTOH 780  
QY 781 TILFNLGWAQAQAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA 840  
DB 781 TILFNLGWAQAQAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA 840  
QY 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAILRRHVGPGGAVQMMNRLTAFASR 900  
DB 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAILRRHVGPGGAVQMMNRLTAFASR 900  
QY 901 GNHVSPTHVPESDAAARVTQLSLTTLQLKRLHQMINEDECSTPCSGSWLRDVMWDTIC 960  
DB 901 GNHVSPTHVPESDAAARVTQLSLTTLQLKRLHQMINEDECSTPCSGSWLRDVMWDTIC 960  
QY 961 TVLTDFKTLQSKLLPRLPGVFFPCQRYKGVWRGDMQITTCGCAQITGHVKNKSMR 1020  
DB 961 TVLTDFKTLQSKLLPRLPGVFFPCQRYKGVWRGDMQITTCGCAQITGHVKNKSMR 1020  
QY 1021 IVGPRTCSTNTHGTPPINAYTTGCTPSPAPNYSRALWRAAEVVEVTRVGDVPHVVTGM 1080  
DB 1021 IVGPRTCSTNTHGTPPINAYTTGCTPSPAPNYSRALWRAAEVVEVTRVGDVPHVVTGM 1080  
QY 1081 TTDNVKCPQVPAPEFTEVDGVRLHRYAPACKPLLRREBVTFLVGLNQYLVGSQLPCEPE 1140  
DB 1081 TTDNVKCPQVPAPEFTEVDGVRLHRYAPACKPLLRREBVTFLVGLNQYLVGSQLPCEPE 1140  
QY 1141 PDVAVLTSLMTPSHITATKRRRLARGSPPSLASSASQLSAPSLKATCTTRHDSFDAD 1200  
DB 1141 PDVAVLTSLMTPSHITATKRRRLARGSPPSLASSASQLSAPSLKATCTTRHDSFDAD 1200  
QY 1201 LIEANLLWQEMGNTTRVESENKVLDSFEPLQAEDEEREVSVPAAELRRSRKPPRAM 1260  
DB 1201 LIEANLLWQEMGNTTRVESENKVLDSFEPLQAEDEEREVSVPAAELRRSRKPPRAM 1260  
QY 1261 PIWARPDPNPPILLESWKDIPVPPVHGCPLPPAKAPPPIPPRRKRTVVLSESTVSALA 1320  
DB 1261 PIWARPDPNPPILLESWKDIPVPPVHGCPLPPAKAPPPIPPRRKRTVVLSESTVSALA 1320  
QY 1321 ELATKTFGSESAVSGTATASPDQSDDDAGSDVESYSSMPLEGECDPDLSDGWS 1380  
DB 1321 ELATKTFGSESAVSGTATASPDQSDDDAGSDVESYSSMPLEGECDPDLSDGWS 1380  
QY 1381 STVSEASEDVCCSMYSYTWGALITPCAEEETKLPIINALSNLLRHNLVYATTSRAS 1440  
DB 1381 STVSEASEDVCCSMYSYTWGALITPCAEEETKLPIINALSNLLRHNLVYATTSRAS 1440  
QY 1441 LRQKVTDFRLQVLDHRYDLVKEMKAKASTVAKLLSVEACKLTTPHARSFKYGAK 1500  
DB 1441 LRQKVTDFRLQVLDHRYDLVKEMKAKASTVAKLLSVEACKLTTPHARSFKYGAK 1500  
QY 1501 DVRLNLSKAVNHRSVWKLLEDTETPIDTITIMAKNEVFCVQPEKGGKPARLVFPDLG 1560

Db 1501 DVRLNSSKAVNHIRSWKDLLEDTETPDTTMAXNEVFCVQPEKGRKPARLIVFPDLG 1560  
 QY 1561 VRVCEKMAIYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFAYDTRCFDS 1620  
 Db 1561 VRVCEKMAIYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMGFAYDTRCFDS 1620  
 QY 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLTNSKGQNGCYRRCRAGSVLT 1680  
 Db 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLTNSKGQNGCYRRCRAGSVLT 1680  
 QY 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
 Db 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
 QY 1741 SAPPGDPPKPYDLELITSCSSNVSAHDASGKRVYLYTRDPTPLAANAETARHTPVN 1800  
 Db 1741 SAPPGDPPKPYDLELITSCSSNVSAHDASGKRVYLYTRDPTPLAANAETARHTPVN 1800  
 QY 1801 SWLGNIIIMYAPTLWARMILMTHFTSILLAQOLEKALDCQIYGACYSLEPDLQIQL 1860  
 Db 1801 SWLGNIIIMYAPTLWARMILMTHFTSILLAQOLEKALDCQIYGACYSLEPDLQIQL 1860  
 QY 1861 HGLSAFSLHSYSGEINRVASCLRLKGLVPPLRVWHRARSVRARLLSQGGAATCGKYL 1920  
 Db 1861 HGLSAFSLHSYSGEINRVASCLRLKGLVPPLRVWHRARSVRARLLSQGGAATCGKYL 1920  
 QY 1921 NWAVRTKLTPIPAASQDLSSWFVAGYSGDDIYHLSLRARPRFWFWCLLLSVGVGIY 1980  
 Db 1921 NWAVRTKLTPIPAASQDLSSWFVAGYSGDDIYHLSLRARPRFWFWCLLLSVGVGIY 1980  
 QY 1981 LLPNR 1985  
 Db 1981 LLPNR 1985

## RESULT 5

ABG30581  
 ID ABG30581 standard; protein; 2201 AA.  
 AC ABG30581;  
 XX  
 DT 21-OCT-2002 (first entry)  
 XX  
 DE Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #1.  
 XX  
 KW Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;  
 XX cell culture replication; NS2/3; NS3/4; NS3; NS5B.  
 OS Hepatitis C virus.  
 XX  
 PN WO200252015-A2.  
 PD 04-JUL-2002.  
 XX  
 PF 20-DEC-2001; 2001WO-CA001843.  
 XX  
 PR 22-DEC-2000; 2000US-0257857P.  
 XX  
 PA (BOEH ) BOEHRINGER INGELHEIM CANADA LTD.  
 XX  
 PI Kukulj G, Pause A;  
 XX  
 DR WPI; 2002-575382/61.  
 DR N-PSDB; ABK88573.  
 XX  
 PT New self-replicating RNA molecules from Hepatitis C virus (HCV), which  
 PT possess enhanced transduction or replication efficiency, useful for  
 PT evaluating potential inhibitors of HCV replication.  
 XX  
 PS Disclosure; Page 49-58; 140pp; English.  
 XX  
 CC The invention describes a self-replicating hepatitis C virus (HCV)

CC polynucleotide molecule comprising a 5'-non translated region (NTR),  
 CC where guanine at position 1 is substituted for adenine, a HCV polyprotein  
 CC region coding for a HCV polyprotein; and a 3'-NTR region. The self-  
 CC replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating  
 CC potential inhibitors of HCV replication. The HCV RNA molecule is also  
 CC useful for efficiently establishing cell culture replication. The self-  
 CC replicating polynucleotide molecule contains a 5'-NTR, where G at  
 CC position 1 is substituted for A, and therefore provides an alternative to  
 CC existing systems comprising a self-replicating HCV RNA molecule that, in  
 CC conjunction with mutations in the HCV non-structural region, such as the  
 CC G(2042)C/R mutations, transduces and/or replicates with greater  
 CC efficiency. This amino acid sequence is encoded by the hepatitis C virus  
 CC replicon APGK12 and contains the viral protease NS2/3, protease complex  
 CC NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B  
 XX  
 SQ Sequence 2201 AA;

Query Match 100.0%; Score 10462; DB 5; Length 2201;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPITAYSQQTRGLGCLITSLTGRDRNOVEGEVQVSTATQSFATCNCVGVWTVHGA 60  
 Db 217 LAPITAYSQQTRGLGCLITSLTGRDRNOVEGEVQVSTATQSFATCNCVGVWTVHGA 276  
 QY 61 GSKTLAGPKPIQMYTNVDQDLVGMQAPPGARSLTCTCGSSDLXVLRHADVIPVRR 120  
 Db 277 GSKTLAGPKPIQMYTNVDQDLVGMQAPPGARSLTCTCGSSDLXVLRHADVIPVRR 336  
 QY 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
 Db 337 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 396  
 QY 181 RSPVETDSSPPAVPOTFOVAHLHAPTGSOKSKTKVPAAYAAQGYKVLVLPNSVAATLPGF 240  
 Db 397 RSPVETDSSPPAVPOTFOVAHLHAPTGSOKSKTKVPAAYAAQGYKVLVLPNSVAATLPGF 456  
 QY 241 AYMSKARGIDPNIRITGVRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 300  
 Db 457 AYMSKARGIDPNIRITGVRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 516  
 QY 301 ILGIGTVLQDAETAGARLVVLATATPPGVTVEHPNIEEVALSSTGEIPEYKAIPIETI 360  
 Db 517 ILGIGTVLQDAETAGARLVVLATATPPGVTVEHPNIEEVALSSTGEIPEYKAIPIETI 576  
 QY 361 KGRHILIFCHSKKKCDLAAKLSGLNNAVAYRGLDVSIVPTSGDVIVVATDALTMTGFT 420  
 Db 577 KGRHILIFCHSKKKCDLAAKLSGLNNAVAYRGLDVSIVPTSGDVIVVATDALTMTGFT 636  
 QY 421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTVQDAVSRSORRGTRGRMGVIYRVTPG 480  
 Db 637 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTVQDAVSRSORRGTRGRMGVIYRVTPG 696  
 QY 481 ERPSGMFDSVLCBVDAGCAWYELTPAETSRLRAYLNTPLGVCQDHLFEWESVFTGL 540  
 Db 697 ERPSGMFDSVLCBVDAGCAWYELTPAETSRLRAYLNTPLGVCQDHLFEWESVFTGL 756  
 QY 541 THIDAHFLSQTQKAGDNPPYLVAQATVCARAQPPPSDQWKKLIRLKPILHGTPL 600  
 Db 757 THIDAHFLSQTQKAGDNPPYLVAQATVCARAQPPPSDQWKKLIRLKPILHGTPL 816  
 QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGSVVIYGR 660  
 Db 817 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGSVVIYGR 876  
 QY 661 IILSKPAILIDREVLVYREFDEMEECASHLPYIEQGMQLAEQFKQKAILGLQTATQAEA 720  
 Db 877 IILSKPAILIDREVLVYREFDEMEECASHLPYIEQGMQLAEQFKQKAILGLQTATQAEA 936  
 QY 721 AAPVVEKWRTLFAFWAKHWNFIISGQYLAGLSTLPGNPAIASIMAFASITSPITTOH 780  
 Db 937 AAPVVEKWRTLFAFWAKHWNFIISGQYLAGLSTLPGNPAIASIMAFASITSPITTOH 996

QY	781	TLFNLGGWVAAQLAPPSAASAFVAGAGIAGAAVSGISGLGVNLDILAGYAGAGALVA	840	Db	2077	HGLSAFSLHSYSPGEINRVASCLRLKGLVPLRVWRHRSVRARLLSQGGRATCGKYL	2136
Db	997	TLFNLGGWVAAQLAPPSAASAFVAGAGIAGAAVSGISGLGVNLDILAGYAGAGALVA	1056	QY	1921	NWAVRTKLKLTIPPAASQDLSNFWAGYSGDDIYHLSLRARPRWFMWCLLLSVGVGIY	1980
QY	841	FKVMSGEMSFEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNRLLIAFASR	900	Db	2137	NWAVRTKLKLTIPPAASQDLSNFWAGYSGDDIYHLSLRARPRWFMWCLLLSVGVGIY	2196
Db	1057	FKVMSGEMSFEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVQVWNRLLIAFASR	1116	QY	1981	LLPNR 1985	
QY	901	GNHVSPTHYVPESDAARVTOILSLTITQLLKRHLQWINEDECSGPCSGWLRDWDWIC	960	Db	2197	LLPNR 2201	
Db	1117	GNHVSPTHYVPESDAARVTOILSLTITQLLKRHLQWINEDECSGPCSGWLRDWDWIC	1176	RESULT 6			
QY	961	TVLDFKTLWLOSKLLPLPGVPFFSCQRYGKVGWREGDGMOTPCGAGIITGHVKNQSGMR	1020	ID	ABG30587	standard; protein; 2201 AA.	
Db	1177	TVLDFKTLWLOSKLLPLPGVPFFSCQRYGKVGWREGDGMOTPCGAGIITGHVKNQSGMR	1236	XX	ABG30587;		
QY	1021	IVGPRCTCNTWGTGTPINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDHYVTGM	1080	XX	21-OCT-2002	(first entry)	
Db	1237	IVGPRCTCNTWGTGTPINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDHYVTGM	1296	XX	Hepatitis C virus NS2/3, NS3/4, NS3 and NS5B #7.		
QY	1081	TTDNVVKCFQVPAPEFFTEVDGVRVLRHYAPACKPLLRBEVPLVGINQYLVGSQPCPE	1140	DE	Self-replicating; hepatitis C virus; HCV; HCV replication inhibitor;		
Db	1297	TTDNVVKCFQVPAPEFFTEVDGVRVLRHYAPACKPLLRBEVPLVGINQYLVGSQPCPE	1356	KW	cell culture replication; NS2/3; NS3/4; NS3; NS5B.		
QY	1141	PDVAVLTSMITDPSHITAETAKRRLARGSPPLSSASSASQLSAPSLKATCTTRHDSPAD	1200	XX	Hepatitis C virus.		
Db	1357	PDVAVLTSMITDPSHITAETAKRRLARGSPPLSSASSASQLSAPSLKATCTTRHDSPAD	1416	XX	WO200252015-A2.		
QY	1201	LLEANLLWRQEMGNITRVESKNVILDSFPLQAEDEDEREVSVPAEILRERSKFPFRAM	1260	XX	04-JUL-2002.		
Db	1417	LLEANLLWRQEMGNITRVESKNVILDSFPLQAEDEDEREVSVPAEILRERSKFPFRAM	1476	XX	20-DEC-2001; 2001WO-CA001843.		
QY	1261	P1WARPDPNPLLESKWDPDYVPPVHVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA	1320	XX	22-DEC-2000; 2000US-0257857P.		
Db	1477	P1WARPDPNPLLESKWDPDYVPPVHVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA	1536	XX	(BOEH ) BOEHRINGER INGELHEIM CANADA LTD.		
QY	1321	ELATKTFGSESAVDSGTATASPDQSDGAGSDVESYSPMPLEGGPDPDLSGWS	1380	XX	Kukolj G, Pause A;		
Db	1537	ELATKTFGSESAVDSGTATASPDQSDGAGSDVESYSPMPLEGGPDPDLSGWS	1596	PI	WPI; 2002-575382/61.		
QY	1381	STVSEASEEDVCCSMYSYTWGALITPCAEEETKLPINALSNSLLRHLNVYATTSRSAS	1440	DR	N-PSDB; ABK88587.		
Db	1597	STVSEASEEDVCCSMYSYTWGALITPCAEEETKLPINALSNSLLRHLNVYATTSRSAS	1656	XX	New self-replicating RNA molecules from Hepatitis C virus (HCV), which		
QY	1441	LROKQVTFDLQVLDHRYDLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYGA	1500	PT	possess enhanced transduction or replication efficiency, useful for		
Db	1657	LROKQVTFDLQVLDHRYDLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYGA	1716	PT	evaluating potential inhibitors of HCV replication.		
QY	1501	DVNLSSKAVNHRSVWVKLLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560	PS	Disclosure; Page 120-129; 140pp; English.		
Db	1717	DVNLSSKAVNHRSVWVKLLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1776	XX	The invention describes a self-replicating hepatitis C virus (HCV)		
QY	1561	VRYCEKMAVDVYSTLPQAVMGSSYGFQYSPGQRFVEFLVNAWAKKCPMGFAYDTRCFDS	1620	CC	polynucleotide molecule comprising a 5'-non translated region (NTR),		
Db	1777	VRYCEKMAVDVYSTLPQAVMGSSYGFQYSPGQRFVEFLVNAWAKKCPMGFAYDTRCFDS	1836	CC	where guanine at position 1 is substituted for adenine, a HCV polypeptide		
QY	1621	TVTENDIRVEESIYOCCLAPARQARISLTERLYIGGPTLNSKGQNGYRCRASGLVT	1680	CC	region coding for a HCV polypeptide; and a 3'-NTR region. The self-		
Db	1837	TVTENDIRVEESIYOCCLAPARQARISLTERLYIGGPTLNSKGQNGYRCRASGLVT	1896	CC	replicating Hepatitis C virus (HCV) RNA molecule is useful for evaluating		
QY	1681	TSCGNTITCYLKAACACRAKLODCTMLVCGDDLAVICESAGTOEDEASLAFTEAMTRY	1740	CC	potential inhibitors of HCV replication. The HCV RNA molecule is also		
Db	1897	TSCGNTITCYLKAACACRAKLODCTMLVCGDDLAVICESAGTOEDEASLAFTEAMTRY	1956	CC	useful for efficiently establishing cell culture replication. The self-		
QY	1741	SAPPDPPKPEYDLELITSCSSNVSVHADSGKRVYVLTDRPTTPLARAWEHTARHTPVN	1800	CC	replicating polynucleotide molecule contains a 5'-NTR, where G at		
Db	1957	SAPPDPPKPEYDLELITSCSSNVSVHADSGKRVYVLTDRPTTPLARAWEHTARHTPVN	2016	CC	position 1 is substituted for A, and therefore provides an alternative to		
QY	1801	SWLGNITMYAPTLMWARMILMTHFFSILLAQOLEKALDCCQYACYSIEPLDLQIQLR	1860	CC	existing systems comprising a self-replicating HCV RNA molecule that, in		
Db	2017	SWLGNITMYAPTLMWARMILMTHFFSILLAQOLEKALDCCQYACYSIEPLDLQIQLR	2076	CC	conjunction with mutations in the HCV non-structural region, such as the		
QY	1861	HGLSAFSLHSYSPGEINRVASCLRLKGLVPLRVWRHRSVRARLLSQGGRATCGKYL	1920	CC	G2042/C/R mutations, transduces and/or replicates with greater		
				CC	efficiency. This amino acid sequence is encoded by the hepatitis C virus		
				CC	replicon APGK12 and contains the viral protease NS2/3, protease complex		
				CC	NS3/4, helicase NS3 and RNA-dependent RNA polymerase NS5B		
				XX	Sequence 2201 AA;		
				QY	Query Match	100.0%; Score 10462; DB 5; Length 2201;	
				Db	Best Local Similarity	99.9%; Pred. No. 0;	
					Matches 1984; Conservative	1; Mismatches	0; Indels 0; Gaps
					1	MAPTAYSQQTGRGLGCIITSLTGRDRNQVEGVEQVYVSTATQSFATCNGVCTVYHGA	60
					217	LAPITAYSQQTGRGLGCIITSLTGRDRNQVEGVEQVYVSTATQSFATCNGVCTVYHGA	276



QY 61 GSKTLGKPGITQMTYTNVDQDLVQWQAPPGARSILTPCTCGSSDLYLUTRHADVIPVRR 120  
DB 277 GSKTLGKPGITQMTYTNVDQDLVQWQAPPGARSILTPCTCGSSDLYLUTRHADVIPVRR 336  
QY 121 GDSRGLSPRPVSYLKGSSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM 180  
DB 337 GDSRGLSPRPVSYLKGSSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM 396  
QY 181 RSPVFTDNSSPAPVQCFQVAVHLHAPGSGSKSTKVPAAAYAGQYKVLVNLNPSVAATLGF 240  
DB 397 RSPVFTDNSSPAPVQCFQVAVHLHAPGSGSKSTKVPAAAYAGQYKVLVNLNPSVAATLGF 456  
QY 241 AYMSKAHGIDNITGVRTITGAPITVSTYKFLADGCGSGGAYDIIICDECHSTDTT 300  
DB 457 AYMSKAHGIDNITGVRTITGAPITVSTYKFLADGCGSGGAYDIIICDECHSTDTT 516  
QY 301 ILGTGVLDOAETAGARLVLATATPPGSVTVPHNIEEVALSSTGEIPFGKAIPIETI 360  
DB 517 ILGTGVLDOAETAGARLVLATATPPGSVTVPHNIEEVALSSTGEIPFGKAIPIETI 576  
QY 361 KGGHLLFCHSKKCDLAAKLSGLGLNAVAYYRGLDVSIVPTSGDVIWATDALMTGFT 420  
DB 577 KGGHLLFCHSKKCDLAAKLSGLGLNAVAYYRGLDVSIVPTSGDVIWATDALMTGFT 636  
QY 421 GPDFSVIDCNTVQTQVDFSLDPTFTTETTVPODAVSRORRGTCGRMGYIRFVTPG 480  
DB 637 GPDFSVIDCNTVQTQVDFSLDPTFTTETTVPODAVSRORRGTCGRMGYIRFVTPG 696  
QY 481 ERPSGMFDSVLCYDAGCAWYELTETVPAETSVRLRAYLNTPLPVCODHLEFVESVFTGL 540  
DB 697 ERPSGMFDSVLCYDAGCAWYELTETVPAETSVRLRAYLNTPLPVCODHLEFVESVFTGL 756  
QY 541 THIDAHLFSLQKAGDNFPLVAYQATVCARAQAPPSWDMKCLIRLXPTLHGPTPL 600  
DB 757 THIDAHLFSLQKAGDNFPLVAYQATVCARAQAPPSWDMKCLIRLXPTLHGPTPL 816  
QY 601 YRLGAVQNEVTHPTITKYIMACMSADLEVVTSTWLVGVGLAALAAAYCLTTGSGVIVGR 660  
DB 817 YRLGAVQNEVTHPTITKYIMACMSADLEVVTSTWLVGVGLAALAAAYCLTTGSGVIVGR 876  
QY 661 IILSGKPAIIPDREVLRYBDEMEECASHLYPIEQGMQLAEQFKQKAIQLLOTATQAEA 720  
DB 877 IILSGKPAIIPDREVLRYBDEMEECASHLYPIEQGMQLAEQFKQKAIQLLOTATQAEA 936  
QY 721 AAPVVEKWRTLLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSLTQ 780  
DB 937 AAPVVEKWRTLLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSLTQ 996  
QY 781 TLLFNILGWVAAQLAPPASAASAFVAGIAGAAVGSIGLGVLDVILAGYAGVAGALVA 840  
DB 997 TLLFNILGWVAAQLAPPASAASAFVAGIAGAAVGSIGLGVLDVILAGYAGVAGALVA 1056  
QY 841 FKVMGEMSTEDLNNLLPAILSPGALVGVVCAAILRRHVGEGEGAVQWNRLLIAFASR 900  
DB 1057 FKVMGEMSTEDLNNLLPAILSPGALVGVVCAAILRRHVGEGEGAVQWNRLLIAFASR 1116  
QY 901 GNVHSPTHVPSDAAARVTOILSSITITQLKRLHWNEDCSTPCSGSWLRDVMWDMIC 960  
DB 1117 GNVHSPTHVPSDAAARVTOILSSITITQLKRLHWNEDCSTPCSGSWLRDVMWDMIC 1176  
QY 961 TVLTDFKTLQSKLLPRLPCVPFFSCQYKGVWRGDGMOTTCPCGAQITGHVKNXGMR 1020  
DB 1177 TVLTDFKTLQSKLLPRLPCVPFFSCQYKGVWRGDGMOTTCPCGAQITGHVKNXGMR 1236  
QY 1021 IVGPRTCSNTHGTPIINAYTTGCTPSPAPNYSRALMRVAEEYEVTRVGDFFHYVTGM 1080  
DB 1237 IVGPRTCSNTHGTPIINAYTTGCTPSPAPNYSRALMRVAEEYEVTRVGDFFHYVTGM 1296  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVRHLRHYPACKPLIREVTVLGNLYVGSQPCPE 1140  
DB 1297 TTDNVKCPQVPAPEFFTEVDGVRHLRHYPACKPLIREVTVLGNLYVGSQPCPE 1356

QY 1141 PDVAVLTSMLTDPGHITAEAKRIARGSPPSLASSSSASQLSKATCTTRHDSPAD 1200  
DB 1357 PDVAVLTSMLTDPGHITAEAKRIARGSPPSLASSSSASQLSKATCTTRHDSPAD 1416  
QY 1201 LIEANLWROQMGNITRVESENKVWILDSPEIQAEBEDEREVSVPABILRRSRKFFPAM 1260  
DB 1417 LIEANLWROQMGNITRVESENKVWILDSPEIQAEBEDEREVSVPABILRRSRKFFPAM 1476  
QY 1261 PIWARPNDYNPLLESWKDPDYPVPHGCPPLPAKAPPIPPRRKRTVVLSSESTVSSALA 1320  
DB 1477 PIWARPNDYNPLLESWKDPDYPVPHGCPPLPAKAPPIPPRRKRTVVLSSESTVSSALA 1536  
QY 1321 ELATKTFGSSSESSAVDSGTATASPDQPSDDGAGSDVESYSSMPLEGECPDPLSDGSW 1380  
DB 1537 ELATKTFGSSSESSAVDSGTATASPDQPSDDGAGSDVESYSSMPLEGECPDPLSDGSW 1596  
QY 1381 STVSESEASDVVCCMSYTTWTGALITPCAABETKLPINALSNSLLRRHNLVYATTSRAS 1440  
DB 1597 STVSESEASDVVCCMSYTTWTGALITPCAABETKLPINALSNSLLRRHNLVYATTSRAS 1656  
QY 1441 LRQKVTFRDLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKEFGYAK 1500  
DB 1657 LRQKVTFRDLQVLDDHYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKEFGYAK 1716  
QY 1501 DVNLSKSAVNHIRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560  
DB 1717 DVNLSKSAVNHIRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKPARLIVFPDLG 1776  
QY 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAYDTRCFDS 1620  
DB 1777 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAYDTRCFDS 1836  
QY 1621 TVTENDIRVEESYIQCCLDAPPEARQAIRSLTERLYTGGPLTNSKGONCGYRRCRASVLT 1880  
DB 1837 TVTENDIRVEESYIQCCLDAPPEARQAIRSLTERLYTGGPLTNSKGONCGYRRCRASVLT 1896  
QY 1681 TSCGNTLTCLYKAAAAACRAAKLQDCTMLVCGDDLVCESAGTQDEDEASLRAFTTEAMTRY 1740  
DB 1897 TSCGNTLTCLYKAAAAACRAAKLQDCTMLVCGDDLVCESAGTQDEDEASLRAFTTEAMTRY 1956  
QY 1741 SAPPGPPPKPEYDLELITSCSSNVSVAHDAAGKRVYLLTRDPTTPLARAWEARHTFPVN 1800  
DB 1957 SAPPGPPPKPEYDLELITSCSSNVSVAHDAAGKRVYLLTRDPTTPLARAWEARHTFPVN 2016  
QY 1801 SWLGNIIIMYAPTILWARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIQRL 1860  
DB 2017 SWLGNIIIMYAPTILWARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIIQRL 2076  
QY 1861 HGLSAPSLHSYSPGEINRVASCLRLKLGVPPLRVWRHARSVRARLLSQGGRATCGKLYF 1920  
DB 2077 HGLSAPSLHSYSPGEINRVASCLRLKLGVPPLRVWRHARSVRARLLSQGGRATCGKLYF 2136  
QY 1921 NNAVRTKLTPIPAASQLDLSWVAGYSGGDIYHLSRARPRFMWCLLLSVGVGIY 1980  
DB 2137 NNAVRTKLTPIPAASQLDLSWVAGYSGGDIYHLSRARPRFMWCLLLSVGVGIY 2196  
QY 1981 LLPNR 1985  
DB 2197 LLPNR 2201

## RESULT 7

ABG32451  
ID ABG32451 standard; protein; 3010 AA.

XX AC ABG32451;

XX AC ABG32451;

DT 15-NOV-2002 (first entry)

XX Hepatitis C virus Con 1 isolate polyprotein.

XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis;

KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;



internal ribosome entry site; IRES; NSSA; HCV replication; polyprotein.  
Hepatitis C virus.  
WO200259321-A2.  
01-AUG-2002.  
16-JAN-2002; 2002WO-EP000526.  
23-JAN-2001; 2001US-0263479P.  
(RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.  
De Francesco R, Migliaccio G, Paonessa G;  
WPI; 2002-599793/64.  
N-PSDB; ABK91411.  
New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
ribosome entry site (IRES) region, useful in studying HCV replication and  
expression.  
Claim 1; Page 34-36; 69pp; English.  
The invention relates to nucleic acid molecules comprising altered HCV  
NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
internal ribosome entry site (IRES) region coding for one or more NS3,  
NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
are detailed in the specification. Also included are (1) an expression  
vector comprising a nucleotide sequence coding for the altered nucleic  
acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
recombinant cell human hepatoma cell comprising the altered nucleic acids  
; (3) a recombinant cell produced by introducing into a human hepatoma  
cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
replicon enhanced cell or which containing a functional HCV replicon; (5)  
an HCV replicon enhanced cells made in the method; and (6) measuring the  
ability of a compound to affect HCV activity. The HCV replication and HCV  
replicon enhanced cells are useful in studying HCV replication and  
expression, and HCV and host cell interactions, producing HCV RNA and  
proteins, and providing a system for measuring the ability of a compound  
to modulate one or more HCV activities e.g. to discover drugs which may  
treat HCV mediated diseases such as liver failure, cirrhosis and  
hepatocellular carcinoma. The present sequence is the HCV replicon Con 1  
polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
and NS5B proteins) used as a basis for the adaptive mutations of the  
invention

1266 AYMSKAHGIDPNIRKGTGRTTITGAPITYSTYVGKFLADGGCGGAYDIIICDECHSTDTT 1325  
301 ILGIGTVLDQAEATAGARLVVLATATPPGVSVPVPHENIEEVALSTGEIPFYGKAIPBETI 360  
1326 ILGIGTVLDQAEATAGARLVVLATATPPGVSVPVPHENIEEVALSTGEIPFYGKAIPBETI 1385  
361 KGRHLIFCHSKKCKDELAALKSLGLNNAVAYYGLDVSVIPTSGDVIIVATDALMTGFT 420  
1386 KGRHLIFCHSKKCKDELAALKSLGLNNAVAYYGLDVSVIPTSGDVIIVATDALMTGFT 1445  
421 GDFSVIDCNVCTQTVDTSFSLDPTFTTETTTVPDQAVSRORRGTGRGMIGYRFTVPG 480  
1446 GDFSVIDCNVCTQTVDTSFSLDPTFTTETTTVPDQAVSRORRGTGRGMIGYRFTVPG 1505  
481 ERPSGMFDSVLCYDAGCAWYELTPAETSVRLRAYLNTPLPVCQDHLFEWESVFTGL 540  
1506 ERPSGMFDSVLCYDAGCAWYELTPAETSVRLRAYLNTPLPVCQDHLFEWESVFTGL 1565  
541 THIDAHFLSOTKQAGDNPPYLVAQATVCARAQAPPPSDQWKKLIRLKPTLHGPTPL 600  
1566 THIDAHFLSOTKQAGDNPPYLVAQATVCARAQAPPPSDQWKKLIRLKPTLHGPTPL 1625  
601 YRLGAVQNEVTTTHPIITKYINACMSADLEVVTSTWLVGGVLAALAAAYCLITGVSIVVGR 660  
1626 YRLGAVQNEVTTTHPIITKYINACMSADLEVVTSTWLVGGVLAALAAAYCLITGVSIVVGR 1685  
661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQALAEQFKAKIIGLOTATQABA 720  
1686 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQALAEQFKAKIIGLOTATQABA 1745  
721 AAPVYESKWRITLAEWAKHWNFISGQYLAGLSTLPGNPAIASMAFTASITSLPTTQH 780  
1746 AAPVYESKWRITLAEWAKHWNFISGQYLAGLSTLPGNPAIASMAFTASITSLPTTQH 1805  
781 TLLFNILGWVAAQALAPPSAAGAFVAGIAGAAVGSIGLGKVLVDILAGYAGVAGALVA 840  
1806 TLLFNILGWVAAQALAPPSAAGAFVAGIAGAAVGSIGLGKVLVDILAGYAGVAGALVA 1865  
841 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVOMNRLIAFASR 900  
1866 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVOMNRLIAFASR 1925  
901 GNVHSPHTVPESDAAARVTOILSSLTITQLKRLHOMINECDSTPCSSWLRDWDWIC 960  
1926 GNVHSPHTVPESDAAARVTOILSSLTITQLKRLHOMINECDSTPCSSWLRDWDWIC 1985  
961 TVLTDFKTLQSKLLPRLPGVPPFFSCQGYKGVWRGDMQITTCPCGAQITGHVKNGSMR 1020  
1986 TVLTDFKTLQSKLLPRLPGVPPFFSCQGYKGVWRGDMQITTCPCGAQITGHVKNGSMR 2045  
1021 IVGPRTCSTNTHGTFPPINAYTTGCTPPSPAPNYSRALWRVAABEYVEVTRVGDHFHVTGM 1080  
2046 IVGPRTCSTNTHGTFPPINAYTTGCTPPSPAPNYSRALWRVAABEYVEVTRVGDHFHVTGM 2105  
1081 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLAREEVTFLVGLNQVLVGSQLPCEPE 1140  
2106 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLAREEVTFLVGLNQVLVGSQLPCEPE 2165  
1141 PDVAVLTSLMTDPSHITAEAKRLARGSPPSLASSASQSLAPSLKATCTTTHDSPAD 1200  
2166 PDVAVLTSLMTDPSHITAEAKRLARGSPPSLASSASQSLAPSLKATCTTTHDSPAD 2225  
1201 LIEANLWRQEMGNITRVESENKVVILDSFEPQAEDEDEREVSVAEILRRSRKPPRAM 1260  
2226 LIEANLWRQEMGNITRVESENKVVILDSFEPQAEDEDEREVSVAEILRRSRKPPRAM 2285  
1261 PIWAPDYNPPLLESKWDPDYVPVVGCPPLPAKAPPIPPRRKRKRWLSESVSSALA 1320  
2286 PIWAPDYNPPLLESKWDPDYVPVVGCPPLPAKAPPIPPRRKRKRWLSESVSSALA 2345  
1321 ELATKTFGSSSESSAVDSGTATASPDQPSDDGAGSDVESYSSMPPLGEFGDPLDSGWS 1380  
2346 ELATKTFGSSSESSAVDSGTATASPDQPSDDGAGSDVESYSSMPPLGEFGDPLDSGWS 2405

Query Match 100.0%; Score 10462; DB 5; Length 3010;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
1 MAPITAYSQOTRGLLCITSLTGRDRNQVEGEVQVSTATQSFATCNGVCWTVYHGA 60  
1026 LAPITAYSQOTRGLLCITSLTGRDRNQVEGEVQVSTATQSFATCNGVCWTVYHGA 1085  
61 GSKTLGAPKGPITOMYTNVDQDLVGWQAPPFARSLTPTCGSSDLYLVTTHADVIPVRRR 120  
1086 GSKTLGAPKGPITOMYTNVDQDLVGWQAPPFARSLTPTCGSSDLYLVTTHADVIPVRRR 1145  
121 GDSRGSLLSPRVSYLKSGGGPLLCPSHGAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 180  
1146 GDSRGSLLSPRVSYLKSGGGPLLCPSHGAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 1205  
181 RSPVFTDSSPAPVQPFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 240  
1206 RSPVFTDSSPAPVQPFQVAHLHAPTGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 1265  
241 AYMSKAHGIDPNIRKGTGRTTITGAPITYSTYVGKFLADGGCGGAYDIIICDECHSTDTT 300

Qy	1381	STVSEASESDVCCSMYSYTWTCALITPCAAEETKLPINALNSLRRHNLVYATTSRSAS	1440
Db	2406	STVSEASESDVCCSMYSYTWTCALITPCAAEETKLPINALNSLRRHNLVYATTSRSAS	2465
Qy	1441	LRQKKVTFPRLOVLDDHYRDVLKWKAKASTVKAKLLSVEEACKLTPPHSARSFYGCAK	1500
Db	2466	LRQKKVTFPRLOVLDDHYRDVLKWKAKASTVKAKLLSVEEACKLTPPHSARSFYGCAK	2525
Qy	1501	DVRNLSSKAVNHRSWKDLLEDTETPDITTIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560
Db	2526	DVRNLSSKAVNHRSWKDLLEDTETPDITTIMAKNEVFCVQPEKGRKPARLIVFPDLG	2585
Qy	1561	VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVFELVNAWKAKCPMGFAIDTRCFDS	1620
Db	2586	VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQRVFELVNAWKAKCPMGFAIDTRCFDS	2645
Qy	1621	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQCGYRRCRAGSVLT	1680
Db	2646	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQCGYRRCRAGSVLT	2705
Qy	1681	TSCGNTLTICYLKAACAAACRAAKLQDCTMLVCGDDLVVICSAGTQDEASLRAFTTEAMTRY	1740
Db	2706	TSCGNTLTICYLKAACAAACRAAKLQDCTMLVCGDDLVVICSAGTQDEASLRAFTTEAMTRY	2765
Qy	1741	SAPPGDPPKPEYDLELITSCSSNSVSAHDASGKRVYLYTRDPTTPLARAAAWETARHTPVN	1800
Db	2766	SAPPGDPPKPEYDLELITSCSSNSVSAHDASGKRVYLYTRDPTTPLARAAAWETARHTPVN	2825
Qy	1801	SWLGNIMVAPTLWARMTLMTFFSILLAQBLEKALDCQIYGACYSIEPLDLPIQIRL	1860
Db	2826	SWLGNIMVAPTLWARMTLMTFFSILLAQBLEKALDCQIYGACYSIEPLDLPIQIRL	2885
Qy	1861	HGLSAFSLHSYSPGEINRVASCLRKLGVPLRVWRHRARSVRARLLSOGGRAATCGKLYF	1920
Db	2886	HGLSAFSLHSYSPGEINRVASCLRKLGVPLRVWRHRARSVRARLLSOGGRAATCGKLYF	2945
Qy	1921	NWAVRTKLTPIPAASQLDSSMFVAGYSGGDIYHLSRARPRPFPMWCLLLLSVGVGTY	1980
Db	2946	NWAVRTKLTPIPAASQLDSSMFVAGYSGGDIYHLSRARPRPFPMWCLLLLSVGVGTY	3005
Qy	1981	LLPNR 1985	
Db	3006	LLPNR 3010	
RESULT 8			
AAE15731			
ID	AAE15731	standard; protein; 1985 AA.	
XX	AC	AAE15731;	
XX	AC	AAE15731;	
DT	12-MAR-2002	(first entry)	
XX			
DE		Hepatitis C virus (HCV) replibBartMan polyprotein variant S1172P.	
XX			
KW		Hepatitis C virus; HCV; transfection; infection; virus neutralisation;	
KW		Gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;	
KW		replibBartMan; mutant; mutein; variant.	
XX			
OS		Hepatitis C virus.	
OS		Synthetic.	
XX			
EH		Key Location/Qualifiers	
FT		Misc-difference 1172	
FT		/note= "Wild type Ser substituted with Pro"	
XX			
PN		WO2001089364-A2.	
XX			
PD		29-NOV-2001.	
XX			
XX		23-MAY-2001; 2001WO-US016822.	
XX			

23-MAY-2000; 2000US-00576989.  
(UNIW ) UNIV WASHINGTON.  
Rice CM, Blight KJ;  
WPI; 2002-066755/09.  
Hepatitis C virus variants having greater transfection efficiency and ability to survive subpassage, useful as a vaccine for immunizing primates to the virus, comprise non-naturally occurring viral sequences.  
Claim 14; Page; 174pp; English.  
The invention relates to Hepatitis C virus (HCV) variants which include polynucleotides comprising non-naturally occurring HCV sequence and HCV variants that have a transfection efficiency and ability to survive subpassage greater than HCV that have wild-type polypeptide coding regions. The polynucleotides of the invention are useful for identifying a cell line that is permissive for infection with HCV and detecting replication of HCV in cells of the cell line. They are also useful for testing a compound for anti-viral properties and for inhibiting HCV infection. They are also useful for the generation of defined HCV virus stocks to develop in vitro and in vivo assays for virus neutralisation, attachment, penetration and entry, structure/function studies on HCV proteins and RNA elements and identification of new antiviral targets, a systematic survey of cell culture variants and conditions to identify, a those that support wild-type and variant HCV RNA replication and particle release, production of adaptive HCV variants capable of more efficiency replication in cell culture, production of HCV variants with altered tissue or species tropism, establishment of alternative animal models for inhibitor evaluation including those supporting HCV variant replication, development of cell-free HCV replication assays, production of immunogenic HCV particles for vaccination, engineering of attenuated HCV derivative as possible vaccine candidates, engineering of attenuated or defective HCV derivatives for expression of heterologous gene products for gene therapy and vaccine applications and for utilisation of the HCV glycoproteins for targeted delivery of therapeutic agents to the liver or other cell types with appropriate receptors. Vaccine comprising these sequences is useful for inducing immunoprotection to HCV in a primate. The present sequence is Hepatitis C virus (HCV) repBartMan polypeptide variant. Note: The present sequence is not shown in the specification but is derived from SEQ ID NO: 3 (AAE1517) shown in page 65 of the specification

SQ	Sequence	1985	AA;
Query Match	100.0%;	Score	10460;
Best Local Similarity	99.9%;	Pred. No.	0;
Matches 1984;	Conservative	0;	Mismatches
		1;	Indels
		0;	Gaps
		0;	
QY	1	MAPITAYSQOTRGLLGCIITSLTGRDRNVEGEVQVVSTATQSFLLATCVNGVCWTVVHGA	60
DB	1	MAPITAYSQOTRGLLGCIITSLTGRDRNVEGEVQVVSTATQSFLLATCVNGVCWTVVHGA	60
QY	61	GSKTLAGPGPITOMYNTVDQDLVHQAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRR	120
DB	61	GSKTLAGPGPITOMYNTVDQDLVHQAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRR	120
QY	121	GDGRGSLSPRPVSYLKGSSGGLPLCPSGHAVGIFRAAAVCTRGVAKAVDPVPVSEMETTM	180
DB	121	GDGRGSLSPRPVSYLKGSSGGLPLCPSGHAVGIFRAAAVCTRGVAKAVDPVPVSEMETTM	180
QY	181	RSPVFTDNSPPAVPQTFQVAHLHAPTGSGKSTKVPAAVAAAGYKVLVLPNPSVAATLGF	240
DB	181	RSPVFTDNSPPAVPQTFQVAHLHAPTGSGKSTKVPAAVAAAGYKVLVLPNPSVAATLGF	240
QY	241	AYMSKAHGIDPNIRTVGRTITTGAPITYSTYKFLADGGCSGGAYDI IICDECHSTDSTT	300
DB	241	AYMSKAHGIDPNIRTVGRTITTGAPITYSTYKFLADGGCSGGAYDI IICDECHSTDSTT	300
QY	301	ILIGITVLDOAETAGARLVWLATATPPGSVTVVPHPNIBEVALSSTGEIPFYGKAIPETI	360

301	ILGIGTVLQDAETAGARLVUATATPPGSSVTVPHPNIEEVALSGTGEIPFYGKAIPETI	360	1441	LROKKVTFDRLQVLDHHDYRDLKEMKAKASTVKAKLLSVEEACKLTTPPHSARSKEGYGAK	1500
361	KGGRHLIFCHSKKCCDELAAGLGLNANAYRGLDVSVIPTSGDVIIVATDALMTGFT	420	1441	LROKKVTFDRLQVLDHHDYRDLKEMKAKASTVKAKLLSVEEACKLTTPPHSARSKEGYGAK	1500
361	KGGRHLIFCHSKKCCDELAAGLGLNANAYRGLDVSVIPTSGDVIIVATDALMTGFT	420	1501	DVRNLSSKANVHRSWKDLLEDTEPIDITIMAKNEVFCVQPEKGGKPKARLIVFPDLG	1560
421	GDPSVIDNCNTCVTQVDFSLDPTFTTETITVPQDAVSRSORRGTRGMRGIVRFVTPG	480	1501	DVRNLSSKANVHRSWKDLLEDTEPIDITIMAKNEVFCVQPEKGGKPKARLIVFPDLG	1560
421	GDPSVIDNCNTCVTQVDFSLDPTFTTETITVPQDAVSRSORRGTRGMRGIVRFVTPG	480	1561	VRVCEKVALYDVVSTLPOAVMGSSYGFQYSPGQRFVFLVNAWKAKKCPMGFAYDTRCFDS	1620
481	BRPSGMFSSVLCCEYDAGCAWYELTPAETSVRILRAYLNTPLPVCQDHLRFESVFTGL	540	1561	VRVCEKVALYDVVSTLPOAVMGSSYGFQYSPGQRFVFLVNAWKAKKCPMGFAYDTRCFDS	1620
481	BRPSGMFSSVLCCEYDAGCAWYELTPAETSVRILRAYLNTPLPVCQDHLRFESVFTGL	540	1621	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGGLTNSKQNGCYRRCRASGVLT	1680
541	THIDAHFISQTKAGDNFPYLVAQATVCARAQAPPPSDQWQKCLIRLKPILHGTPTLL	600	1621	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGGLTNSKQNGCYRRCRASGVLT	1680
541	THIDAHFISQTKAGDNFPYLVAQATVCARAQAPPPSDQWQKCLIRLKPILHGTPTLL	600	1681	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY	1740
601	YRLGAVQNEVTTTHPITIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTVVIVGR	660	1681	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY	1740
601	YRLGAVQNEVTTTHPITIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTVVIVGR	660	1741	SAPPGDPKPEVDLELITSCSSNVSVAHDAAGKRVYVYLTTRDPTTPLARAAMETARHTPVN	1800
661	IILSGKPAIIPREVLYREFDEMEECASHLPYIEQWQLAEQFKQKALGLLQATKQAEA	720	1741	SAPPGDPKPEVDLELITSCSSNVSVAHDAAGKRVYVYLTTRDPTTPLARAAMETARHTPVN	1800
661	IILSGKPAIIPREVLYREFDEMEECASHLPYIEQWQLAEQFKQKALGLLQATKQAEA	720	1801	SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPDLPIQL	1860
721	AAPVVEKWRITLFAFWAKHWNFIISGIVLAGLSTLPGNPATASLMAFTASITSLTTOH	780	1801	SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPDLPIQL	1860
721	AAPVVEKWRITLFAFWAKHWNFIISGIVLAGLSTLPGNPATASLMAFTASITSLTTOH	780	1861	HGLSAFSLHSYSPGINRVASCLRGVPPPLVWRHRSVRARLLSQGGRATCGKYL	1920
781	TLILFNILGWAAQLAPPSSAFAVAGAGIAGAVSGIGLKVLDILAGYAGVAGALVA	840	1861	HGLSAFSLHSYSPGINRVASCLRGVPPPLVWRHRSVRARLLSQGGRATCGKYL	1920
781	TLILFNILGWAAQLAPPSSAFAVAGAGIAGAVSGIGLKVLDILAGYAGVAGALVA	840	1921	NWAVRTKLTPIPAASQDLSSWVAGVSGGDIYHLSRPRFWMFMCLLLSVGVGIY	1980
841	EKVMGSEMPSTEDLNLLPAIISPGALVGVVCAAILRRHVGGEGAVQWMLRIAFASR	900	1921	NWAVRTKLTPIPAASQDLSSWVAGVSGGDIYHLSRPRFWMFMCLLLSVGVGIY	1980
841	EKVMGSEMPSTEDLNLLPAIISPGALVGVVCAAILRRHVGGEGAVQWMLRIAFASR	900	1981	LLPNR 1985	
901	GNHVSPTHVPESDAAARVQILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC	960	1981	LLPNR 1985	
901	GNHVSPTHVPESDAAARVQILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC	960	RESULT 9		
961	TVLTDFTKWLQSKLLPRLPVPFSCORGKVGWVGDMQTTCPGQAQITGHVKXGSMR	1020	AAE15730	standard; protein; 1985 AA.	
961	TVLTDFTKWLQSKLLPRLPVPFSCORGKVGWVGDMQTTCPGQAQITGHVKXGSMR	1020	ID	AAE15730	
1021	IVGPRICSNTHGTFPPINATYTGCTPSPAPNYSRALRWVAABEYVEVTRVGDVHYVTGM	1080	XX	AAE15730;	
1021	IVGPRICSNTHGTFPPINATYTGCTPSPAPNYSRALRWVAABEYVEVTRVGDVHYVTGM	1080	AC		
1081	TTDNVKKPCQVPAPEFTEVDGVRHLRYAPACKPDLREEVTFVLGNQYLVGSQLPCEPE	1140	XX	12-MAR-2002 (first entry)	
1081	TTDNVKKPCQVPAPEFTEVDGVRHLRYAPACKPDLREEVTFVLGNQYLVGSQLPCEPE	1140	DT		
1141	PDVAULTSMLTDPHSHTAETAKRELARGSPPLASSASQSLASPLKATCTTRHDSPAD	1200	XX	Hepatitis C virus (HCV) replibBartMan polyprotein variant S1172C.	
1141	PDVAULTSMLTDPHSHTAETAKRELARGSPPLASSASQSLASPLKATCTTRHDSPAD	1200	DE		
1201	LIEANLLWRQBMGNGNIIRVSENVKWLIDSPFELQAEDEDEREVSVPAEILRRSKFFPRAM	1260	XX	Hepatitis C virus; HCV; transfection; infection; virus neutralisation;	
1201	LIEANLLWRQBMGNGNIIRVSENVKWLIDSPFELQAEDEDEREVSVPAEILRRSKFFPRAM	1260	KW	gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;	
1261	PIWARPDPNPLLESWKDPDVPVPHGCPILPPAKAPPIPPPRKRTVILSESTVSSALA	1320	KM	replibBartMan; mutant; mutein; variant.	
1261	PIWARPDPNPLLESWKDPDVPVPHGCPILPPAKAPPIPPPRKRTVILSESTVSSALA	1320	XX	Hepatitis C virus.	
1321	ELATKTFGSSSAVDSTATASDPQDDGAGSDVESYSSMPPEGECDPDLSDGSW	1380	OS	Synthetic.	
1321	ELATKTFGSSSAVDSTATASDPQDDGAGSDVESYSSMPPEGECDPDLSDGSW	1380	XX	Key	
1381	STVSEASEDDVCCSMSTYTWGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS	1440	PH	Location/Qualifiers	
1381	STVSEASEDDVCCSMSTYTWGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS	1440	FT	Misc-difference 1172	
			FT	/note= "Wild type Ser substituted with Cys"	
			XX	WO200189364-A2.	
			XX	29-NOV-2001.	
			XX	23-MAY-2001; 2001WO-US016822.	
			XX	23-MAY-2000; 2000US-00576989.	
			XX	(UNIW ) UNIV WASHINGTON.	

PI Rice CM, Blight KJ;  
 DR WPI: 2002-066755/09.  
 XX  
 XX Hepatitis C virus variants having greater transfection efficiency and  
 PT ability to survive subpassage, useful as a vaccine for immunizing primate  
 PT to the virus, comprise non-naturally occurring viral sequences.  
 XX  
 XX Claim 14; Page; 174pp; English.  
 XX  
 CC The invention relates to Hepatitis C virus (HCV) variants which include  
 CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
 CC variants that have a transfection efficiency and ability to survive  
 CC subpassage greater than HCV that have wild-type polypeptide coding  
 CC regions. The polynucleotides of the invention are useful for identifying  
 CC a cell line that is permissive for infection with HCV and detecting  
 CC replication of HCV in cells of the cell line. They are also useful for  
 CC testing a compound for anti-viral properties and for inhibiting HCV  
 CC infection. They are also useful for the generation of defined HCV virus  
 CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
 CC attachment, penetration and entry, structure/function studies on HCV  
 CC proteins and RNA elements and identification of new antiviral targets, a  
 CC systematic survey of cell culture systems and conditions to identify  
 CC those that support wild-type and variant HCV RNA replication and particle  
 CC release, production of adaptive HCV variants capable of more efficiency  
 CC replication in cell culture, production of HCV variants with altered  
 CC tissue or species tropism, establishment of alternative animal models for  
 CC inhibitor evaluation including those supporting HCV variant replication,  
 CC development of cell-free HCV replication assays, production of  
 CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
 CC derivatives as possible vaccine candidates, engineering of attenuated or  
 CC defective HCV derivatives for expression of heterologous gene products  
 CC for gene therapy and vaccine applications and for utilisation of the HCV  
 CC glycoproteins for targeted delivery of therapeutic agents to the liver  
 CC or other cell types with appropriate receptors. Vaccine comprising these  
 CC sequences is useful for inducing immunoprotection to HCV in a primate.  
 CC The present sequence is Hepatitis C virus (HCV) repBartMan polypeptide  
 CC variant. Note: The present sequence is not shown in the specification but  
 CC is derived from SEQ ID NO: 3 (AAE15171) shown in page 65 of the  
 CC specification  
 XX  
 SQ Sequence 1985 AA;  
 Query Match 100.0%; Score 10460; DB 5; Length 1985;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPITAYSQOTRGLLCIITSLTGRDRNQVEGVVQVSTATQSLATCVNGVCWTVYHGA 60  
 DB 1 MAPITAYSQOTRGLLCIITSLTGRDRNQVEGVVQVSTATQSLATCVNGVCWTVYHGA 60  
 QY 61 GSKTLGAPKGPITOMYTNVDQVLGVQWAPPGARSLTPTCTGSSDLVLTTHADVIPVRRR 120  
 DB 61 GSKTLGAPKGPITOMYTNVDQVLGVQWAPPGARSLTPTCTGSSDLVLTTHADVIPVRRR 120  
 QY 121 GDSRGLLSPPRVSYLKGSGGGLLCPGSHAVGIFRAAVCTRCVAKADVFPVSEMETTM 180  
 DB 121 GDSRGLLSPPRVSYLKGSGGGLLCPGSHAVGIFRAAVCTRCVAKADVFPVSEMETTM 180  
 QY 181 RSPVFTDSSPPAVPOTFOVAHLHAPTGSKSTKVPAAVAAQYKVLVLPNSVAATLGF 240  
 DB 181 RSPVFTDSSPPAVPOTFOVAHLHAPTGSKSTKVPAAVAAQYKVLVLPNSVAATLGF 240  
 QY 241 AYMSKAHGIDPNIRTCVRIITCAPITYTYGKFLADGGCGGAYDIIICDECHSTDSTT 300  
 DB 241 AYMSKAHGIDPNIRTCVRIITCAPITYTYGKFLADGGCGGAYDIIICDECHSTDSTT 300  
 QY 301 ILGIGTVLQDQATAGARLVVLTATPPGVTVPHPNI E VALSSSTGEIPFYKALPIETI 360  
 DB 301 ILGIGTVLQDQATAGARLVVLTATPPGVTVPHPNI E VALSSSTGEIPFYKALPIETI 360  
 QY 361 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGDLVSVIPTSGDVIIVVATDALMTGFT 420

Db 361 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGDLVSVIPTSGDVIIVVATDALMTGFT 420  
 QY 421 GDFDSVIDCNTCVTQTVDFSLDPTFTTETTTVPDAVSRQRRGTGRMGIVYFVTPG 480  
 Db 421 GDFDSVIDCNTCVTQTVDFSLDPTFTTETTTVPDAVSRQRRGTGRMGIVYFVTPG 480  
 QY 481 ERPSGMDSSVLCBCEYDAGCAWYELTTPAETSVRRLAYLNTFGLPVCQDHLFWESVFTGL 540  
 Db 481 ERPSGMDSSVLCBCEYDAGCAWYELTTPAETSVRRLAYLNTFGLPVCQDHLFWESVFTGL 540  
 QY 541 THIDAHFLSOTKQAGDNFFYLVAOATVCARAQAPPSWDQWKLRLKLTGLHGTPTLL 600  
 Db 541 THIDAHFLSOTKQAGDNFFYLVAOATVCARAQAPPSWDQWKLRLKLTGLHGTPTLL 600  
 QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTGVIIVGR 660  
 Db 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTGVIIVGR 660  
 QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQPKQKAIIGLLQTATKQAE 720  
 Db 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQPKQKAIIGLLQTATKQAE 720  
 QY 721 AAPVSEKWTLEAFWAKHWNFIISGLIYLAGLSTLPGNPAIASLMAFTASITSLTQH 780  
 Db 721 AAPVSEKWTLEAFWAKHWNFIISGLIYLAGLSTLPGNPAIASLMAFTASITSLTQH 780  
 QY 781 TLLFNILGGVAAQALAPPSAASAFVAGIAGAAVGSIGLGKVLVDILLAGYAGVAGALVA 840  
 Db 781 TLLFNILGGVAAQALAPPSAASAFVAGIAGAAVGSIGLGKVLVDILLAGYAGVAGALVA 840  
 QY 841 FKVMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEVAVQWNNRLIAFASR 900  
 Db 841 FKVMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEVAVQWNNRLIAFASR 900  
 QY 901 GNVSPHYVPESDAARVQTILSSLTITQLKRLHWNEDCSTPCSGSWLDRVNDWIC 960  
 Db 901 GNVSPHYVPESDAARVQTILSSLTITQLKRLHWNEDCSTPCSGSWLDRVNDWIC 960  
 QY 961 TVLTDFKTLQSKLLPRLPGVPFFSCQYGVWVRGDMQTTCPGCAQITGHVKGSMR 1020  
 Db 961 TVLTDFKTLQSKLLPRLPGVPFFSCQYGVWVRGDMQTTCPGCAQITGHVKGSMR 1020  
 QY 1021 IVGPRTCSNTWHTGTFPINAYTTGCTPSPAPNYSRALRVAAEYVEVTRVGDVHYVTGM 1080  
 Db 1021 IVGPRTCSNTWHTGTFPINAYTTGCTPSPAPNYSRALRVAAEYVEVTRVGDVHYVTGM 1080  
 QY 1081 TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKELLREEVTLVGLNQYLVGSQLPCEPE 1140  
 Db 1081 TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKELLREEVTLVGLNQYLVGSQLPCEPE 1140  
 QY 1141 PDVAVLTSMLTDPSSHITAEAKRLARGSPPSLASSASSQLSAPSLKATCTTRHDSFDAD 1200  
 Db 1141 PDVAVLTSMLTDPSSHITAEAKRLARGSPPSLASSASSQLSAPSLKATCTTRHDSFDAD 1200  
 QY 1201 LIEANLLWRQEMGNITRVESKNNVILDSFPELQAEDEREVSPAEILRRSRKFPFRAM 1260  
 Db 1201 LIEANLLWRQEMGNITRVESKNNVILDSFPELQAEDEREVSPAEILRRSRKFPFRAM 1260  
 QY 1261 PWARPDYNNPILLESWKOPDYVPPVHGCPLPAKAPPIPPRRKRKTVLSESTVSSALA 1320  
 Db 1261 PWARPDYNNPILLESWKOPDYVPPVHGCPLPAKAPPIPPRRKRKTVLSESTVSSALA 1320  
 QY 1321 ELATKTFGSSSAVDSTATASPPQSDGSDGSDVESYSSMPPLEGEPPGDDSDGSM 1380  
 Db 1321 ELATKTFGSSSAVDSTATASPPQSDGSDGSDVESYSSMPPLEGEPPGDDSDGSM 1380  
 QY 1381 STVSEASEDVVCCSMSTWTGALTTPCAAEETKLPINALSNLRLHNLVYATTSRSAS 1440  
 Db 1381 STVSEASEDVVCCSMSTWTGALTTPCAAEETKLPINALSNLRLHNLVYATTSRSAS 1440  
 QY 1441 LRQKVTVDRLQVLDHVDYRDLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFYGAK 1500  
 Db 1441 LRQKVTVDRLQVLDHVDYRDLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFYGAK 1500



Db 421 GDFSDVIDCNTVCTQVDFSLDPTETITTTVPDAVSQRGRGTGRGRMCIYRFVTGP 480  
QY 481 ERPSGMDSSVLCYCDAGCAWYELTPAETSURLPAYLNTPLGCPVQDHLFPWESVFTGL 540  
Db 481 ERPSGMDSSVLCYCDAGCAWYELTPAETSURLPAYLNTPLGCPVQDHLFPWESVFTGL 540  
QY 541 THIDAHFLSQTQKAGDNFFYLVAQATVCARAQAPPPSWDQWKLIRLKPITLHGPTPL 600  
Db 541 THIDAHFLSQTQKAGDNFFYLVAQATVCARAQAPPPSWDQWKLIRLKPITLHGPTPL 600  
QY 601 YRLGAVQNEVTTTPIKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGTSVIVGR 660  
Db 601 YRLGAVQNEVTTTPIKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGTSVIVGR 660  
QY 661 IILSGKPAIIPDRVLVREFDEMECEASHLPIYEQGQOLAFQFKQKAIIGLLOTKAQEA 720  
Db 661 IILSGKPAIIPDRVLVREFDEMECEASHLPIYEQGQOLAFQFKQKAIIGLLOTKAQEA 720  
QY 721 AAPVVEKWTLEAFWAKHMWNFIISGQYLAGLSTLPGNPAIASLMAFTASITSPLOTQ 780  
Db 721 AAPVVEKWTLEAFWAKHMWNFIISGQYLAGLSTLPGNPAIASLMAFTASITSPLOTQ 780  
QY 781 TLLFNILGGVAAQALAPPSAASAPVGAAGVAGSISGLGKVLVDIILAGYGAGVAGALVA 840  
Db 781 TLLFNILGGVAAQALAPPSAASAPVGAAGVAGSISGLGKVLVDIILAGYGAGVAGALVA 840  
QY 841 FKVMSGEMPTEDIVNLPAIIPALPGALVGVVCAAILRRHVGPGEAVQWNRLLIAPASR 900  
Db 841 FKVMSGEMPTEDIVNLPAIIPALPGALVGVVCAAILRRHVGPGEAVQWNRLLIAPASR 900  
QY 901 GNHVSPTHYPVPSDAAARVTQIILSSLTITQILKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
Db 901 GNHVSPTHYPVPSDAAARVTQIILSSLTITQILKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
QY 961 TVLTDFKTLQSKLLPRLPGVPFFSCORGKYGVMRGDGMQTTCPGCAQITGHVKGSMR 1020  
Db 961 TVLTDFKTLQSKLLPRLPGVPFFSCORGKYGVMRGDGMQTTCPGCAQITGHVKGSMR 1020  
QY 1021 IVGPRTCSNTHGTFPPINATTCGCTPSPAPNYSRALWRVAEEYVEVTRVGFHVTGM 1080  
Db 1021 IVGPRTCSNTHGTFPPINATTCGCTPSPAPNYSRALWRVAEEYVEVTRVGFHVTGM 1080  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPLREVEVTLVGLNOYLNGSQLPCEPE 1140  
Db 1081 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPLREVEVTLVGLNOYLNGSQLPCEPE 1140  
QY 1141 PDVAVLTSLMTDPSHTATATKRLARGSPPSLIASSASQLSAPSLKATCTTRHDSFAD 1200  
Db 1141 PDVAVLTSLMTDPSHTATATKRLARGSPPSLIASSASQLSAPSLKATCTTRHDSFAD 1200  
QY 1201 LIEANLLWRQEGNITRVESENKVVILDSFPLQAEDEDEREVSVPAILRRSRKPPRAM 1260  
Db 1201 LIEANLLWRQEGNITRVESENKVVILDSFPLQAEDEDEREVSVPAILRRSRKPPRAM 1260  
QY 1261 PIWARPDPNPPLESKWDYDYPVPHGCPPLPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 1261 PIWARPDPNPPLESKWDYDYPVPHGCPPLPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
QY 1321 ELATKTFGSSSAVDSGTATASPOQSDDGAGSDVESYSSMPLEGEPPDLSGDSW 1380  
Db 1321 ELATKTFGSSSAVDSGTATASPOQSDDGAGSDVESYSSMPLEGEPPDLSGDSW 1380  
QY 1381 STVSEASDDVCCSMYSYTWTCALITPCAAETKLPINALSNLIRHNLVYATTSRGAS 1440  
Db 1381 STVSEASDDVCCSMYSYTWTCALITPCAAETKLPINALSNLIRHNLVYATTSRGAS 1440  
QY 1441 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKEGYGAK 1500  
Db 1441 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKEGYGAK 1500  
QY 1501 DVRLNLSKAVNHRSWKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
Db 1501 DVRLNLSKAVNHRSWKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560

QY 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSGQRFVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
Db 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSGQRFVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
QY 1621 TVTENDIRVESIYQCCDLAPAPARQAIIRSLTERLYIGGPLTNSKGQNGYRRCRASGVLT 1680  
Db 1621 TVTENDIRVESIYQCCDLAPAPARQAIIRSLTERLYIGGPLTNSKGQNGYRRCRASGVLT 1680  
QY 1681 TSCGNTLTCYLKAAACRAAKLOCTMLVCGDDLVVICESAGTOEDASLAFTEAMTRY 1740  
Db 1681 TSCGNTLTCYLKAAACRAAKLOCTMLVCGDDLVVICESAGTOEDASLAFTEAMTRY 1740  
QY 1741 SAPPGDPKPEYDLELITSCSSNVSVADHAGSKRVYVLTDRPTTPLARAAMETARHTFVN 1800  
Db 1741 SAPPGDPKPEYDLELITSCSSNVSVADHAGSKRVYVLTDRPTTPLARAAMETARHTFVN 1800  
QY 1801 SWLGNIIIMYAPTILWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPIQL 1860  
Db 1801 SWLGNIIIMYAPTILWARMILMTHFFSILLAQEQLEKALDCQIYGACYSIEPLDLPIQL 1860  
QY 1861 HGLSAPSLHSYSPEINRVASCLRKLGVPPLRVWRHARSVRARLLSQGGRAATCGKYL 1920  
Db 1861 HGLSAPSLHSYSPEINRVASCLRKLGVPPLRVWRHARSVRARLLSQGGRAATCGKYL 1920  
QY 1921 NWAVRTKLTPIPAASQLDLSWFWAGYSGGDIYHLSLRARPRFWMCCLLLSVGVGIY 1980  
Db 1921 NWAVRTKLTPIPAASQLDLSWFWAGYSGGDIYHLSLRARPRFWMCCLLLSVGVGIY 1980  
QY 1981 LLENR 1985  
Db 1981 LLENR 1985  
RESULT 11  
AAE15728  
ID AAE15728 standard; protein; 1985 AA.  
XX  
AC AAE15728;  
XX  
DT 12-MAR-2002 (first entry)  
XX  
DE Hepatitis C virus (HCV) replibBartMan polyprotein variant R1164G.  
XX  
KW Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
KW replibBartMan; mutant; mutein; variant.  
XX  
OS Hepatitis C virus.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FT Misc-difference 1164 /note= "wild type Arg substituted with Gly"  
XX  
XX WO200189364-A2.  
XX  
XX 29-NOV-2001.  
XX  
XX 23-MAY-2001; 2001WO-US016822.  
XX  
XX 23-MAY-2000; 2000US-00576989.  
XX (UNIW ) UNIV WASHINGTON.  
XX  
XX Rice CM, Blight KJ;  
XX  
XX WPI; 2002-066755/09.  
XX  
XX Hepatitis C virus variants having greater transfection efficiency and  
XX ability to survive subpassage, useful as a vaccine for immunizing primate  
XX to the virus, comprise non-naturally occurring viral sequences.  
XX



PS	Claim 14; Page; 174pp; English.	
XX	The invention relates to Hepatitis C virus (HCV) variants which include	
CC	polynucleotides comprising non-naturally occurring HCV sequence and HCV	
CC	variants that have a transfection efficiency and ability to survive	
CC	subpassage greater than HCV that have wild-type polypeptide coding	
CC	regions. The polynucleotides of the invention are useful for identifying	
CC	a cell line that is permissive for infection with HCV and detecting	
CC	replication of HCV in cells of the cell line. They are also useful for	
CC	testing a compound for anti-viral properties and for inhibiting HCV	
CC	infection. They are also useful for the generation of defined HCV virus	
CC	stocks to develop in vitro and in vivo assays for virus neutralisation,	
CC	attachment, penetration and entry, structure/function studies on HCV	
CC	proteins and RNA elements and identification of new antiviral targets, a	
CC	systematic survey of cell culture systems and conditions to identify	
CC	those that support wild-type and variant HCV RNA replication and particle	
CC	release, production of adaptive HCV variants capable of more efficiency	
CC	replication in cell culture, production of HCV variants with altered	
CC	tissue or species tropism, establishment of alternative animal models for	
CC	inhibitor evaluation including those supporting HCV variant replication,	
CC	development of cell-free HCV replication assays, production of	
CC	immunogenic HCV particles for vaccination, engineering of attenuated HCV	
CC	derivatives as possible vaccine candidates, engineering of attenuated or	
CC	defective HCV derivatives for expression of heterologous gene products	
CC	for gene therapy and vaccine applications and for utilisation of the HCV	
CC	glycoproteins for targeted delivery of therapeutic agents to the liver	
CC	or other cell types with appropriate receptors. Vaccine comprising these	
CC	sequences is useful for inducing immunoprotection to HCV in a primate.	
CC	The present sequence is Hepatitis C virus (HCV) replBartMan polypeptide	
CC	variant. Note: The present sequence is not shown in the specification but	
CC	is derived from SEQ ID NO: 3 (AAE15717) shown in page 65 of the	
CC	specification	
XX		
SQ	Sequence 1985 AA;	
	Query Match 99.9%; Score 10458; DB 5; Length 1985;	
	Best Local Similarity 99.9%; Pred. No. 0;	
	Matches 1984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1 MAPITAYSQOTRGLLCIIITSLTGRDRNVEGEVQVWSTATQSLATCNGVCWTVYHGA 60	
DB	1 MAPITAYSQOTRGLLCIIITSLTGRDRNVEGEVQVWSTATQSLATCNGVCWTVYHGA 60	
QY	61 GSKTLGPKGPIOMYTNVDQDLVGCAPPGASLTPTCGSDLYLVTRHADVIPIVRR 120	
DB	61 GSKTLGPKGPIOMYTNVDQDLVGCAPPGASLTPTCGSDLYLVTRHADVIPIVRR 120	
QY	121 GDSRGLLPSPVSYLKGSGGGLCPGSHAVGIFRAAVCTRGVAKAVDFPVVSMETTM 180	
DB	121 GDSRGLLPSPVSYLKGSGGGLCPGSHAVGIFRAAVCTRGVAKAVDFPVVSMETTM 180	
QY	181 RSPVFTDNSPPAVPQTFQVAHLHAPGSGKSTKVPAAVAAQGYKVLVLPNSVAAILGFG 240	
DB	181 RSPVFTDNSPPAVPQTFQVAHLHAPGSGKSTKVPAAVAAQGYKVLVLPNSVAAILGFG 240	
QY	241 AYMSKAGHDIPNRTGVRTITTCAPITYSYGKFLADGGCGGAYDIIICDECHSTDSTT 300	
DB	241 AYMSKAGHDIPNRTGVRTITTCAPITYSYGKFLADGGCGGAYDIIICDECHSTDSTT 300	
QY	301 ILGIGTVLDAQETAGARLVVLTATPGSVTVVPHNIEEVALSGTGPIFYGAIPETI 360	
DB	301 ILGIGTVLDAQETAGARLVVLTATPGSVTVVPHNIEEVALSGTGPIFYGAIPETI 360	
QY	361 KGRHILFCHSKKKCDLAKLSGLGLNNAVAYRGLDVSVIPSGDVIVVATDALMTGFT 420	
DB	361 KGRHILFCHSKKKCDLAKLSGLGLNNAVAYRGLDVSVIPSGDVIVVATDALMTGFT 420	
QY	421 GDPDSVLDKNTCYTQTQWDFSLDPTFTTETTPQDAVRSORGRCTGRGMGIYRFVTPG 480	
DB	421 GDPDSVLDKNTCYTQTQWDFSLDPTFTTETTPQDAVRSORGRCTGRGMGIYRFVTPG 480	
QY	481 ERPSGMFDSSVLCBCEYDAGCAWYELTTPAETSVRLRAYLNTPLPVCODHLEFWESVFTGL 540	
DB	481 ERPSGMFDSSVLCBCEYDAGCAWYELTTPAETSVRLRAYLNTPLPVCODHLEFWESVFTGL 540	
DB	481 ERPSGMFDSSVLCBCEYDAGCAWYELTTPAETSVRLRAYLNTPLPVCODHLEFWESVFTGL 540	
QY	541 THIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPPSWDMWKCLIRLKPDLHGPTPL 600	
DB	541 THIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPPSWDMWKCLIRLKPDLHGPTPL 600	
QY	601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVSTTWLVGGVLAALAAAYCLTTGSGVIVGR 660	
DB	601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVSTTWLVGGVLAALAAAYCLTTGSGVIVGR 660	
QY	661 IILSGKPAIPDRVLVREFDEMEECASHLPIYISQWQLAEQFKOKAIGLLQATKQAEA 720	
DB	661 IILSGKPAIPDRVLVREFDEMEECASHLPIYISQWQLAEQFKOKAIGLLQATKQAEA 720	
QY	721 AAPVVEKWRLEAFWAKHWNFIISQYLAGLSTLPGNPAIASIMAFASITSLPTTQH 780	
DB	721 AAPVVEKWRLEAFWAKHWNFIISQYLAGLSTLPGNPAIASIMAFASITSLPTTQH 780	
QY	781 TLLFNILGGWAAQIAPPASAASAFVAGIAGAAVSGISGLKVLDILAGYGAGVAGALVA 840	
DB	781 TLLFNILGGWAAQIAPPASAASAFVAGIAGAAVSGISGLKVLDILAGYGAGVAGALVA 840	
QY	841 PKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPBGAGVQWNNRLIAFASR 900	
DB	841 PKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPBGAGVQWNNRLIAFASR 900	
QY	901 GNHVSPTHVSPESDAAARVTQILSLTITQLKRLHQMINEPCSTPCSGWLRDWDWTC 960	
DB	901 GNHVSPTHVSPESDAAARVTQILSLTITQLKRLHQMINEPCSTPCSGWLRDWDWTC 960	
QY	961 TVLTDFTKWLQSKLLPRLPGVPPFSCQRYGKVGWGDGIMQTTCPGQAQITGHVKNXGMR 1020	
DB	961 TVLTDFTKWLQSKLLPRLPGVPPFSCQRYGKVGWGDGIMQTTCPGQAQITGHVKNXGMR 1020	
QY	1021 IVGPRTCSNTWHGTPIINAYTTGCTPSPAPNYSRALMRVAEEVEVTRVGDHYVTGM 1080	
DB	1021 IVGPRTCSNTWHGTPIINAYTTGCTPSPAPNYSRALMRVAEEVEVTRVGDHYVTGM 1080	
QY	1081 TTDNVKCPQVPAPFFTEVDGVRHRYAPACKPLLEEVTVLGLNQVLVGSQLPCEPE 1140	
DB	1081 TTDNVKCPQVPAPFFTEVDGVRHRYAPACKPLLEEVTVLGLNQVLVGSQLPCEPE 1140	
QY	1141 PDVAVLTSMLTDPHSHTAETAKRGLARGSPPSLAASSASQLSAPSKATCTTRHDSPPAD 1200	
DB	1141 PDVAVLTSMLTDPHSHTAETAKRGLARGSPPSLAASSASQLSAPSKATCTTRHDSPPAD 1200	
QY	1201 LJEANLLMQEOMGNNITRVESKVVILDSFPLOAEDEDEREVSVAEILRRSRKFFRAM 1260	
DB	1201 LJEANLLMQEOMGNNITRVESKVVILDSFPLOAEDEDEREVSVAEILRRSRKFFRAM 1260	
QY	1261 PIWARPDPNPPILLESWKDPDYPPVHVGCPLPPAKAPPIPPRRKRKTVLSESTVSALA 1320	
DB	1261 PIWARPDPNPPILLESWKDPDYPPVHVGCPLPPAKAPPIPPRRKRKTVLSESTVSALA 1320	
QY	1321 ELAKTFTGSSSAVDSGTATASPDQSDDGAGSDVESYSSMPLEGEDEPDLDGSGW 1380	
DB	1321 ELAKTFTGSSSAVDSGTATASPDQSDDGAGSDVESYSSMPLEGEDEPDLDGSGW 1380	
QY	1381 STVSEBASEDVVCCSMYTWGTALITPCAEEETKLPINALSNLLRHNLVYATTSSRSAS 1440	
DB	1381 STVSEBASEDVVCCSMYTWGTALITPCAEEETKLPINALSNLLRHNLVYATTSSRSAS 1440	
QY	1441 LRQKVTFFORLQVLDHRYDLVKEMKAKASTVAKILSVEEACKLTPPHSARSKFGYGAK 1500	
DB	1441 LRQKVTFFORLQVLDHRYDLVKEMKAKASTVAKILSVEEACKLTPPHSARSKFGYGAK 1500	
QY	1501 DVNRLSKANVHRSVWKDLLEDETEPIDTTIIIAKNEVFCVQPEKGRKPARLIVFPDLG 1560	
DB	1501 DVNRLSKANVHRSVWKDLLEDETEPIDTTIIIAKNEVFCVQPEKGRKPARLIVFPDLG 1560	
QY	1561 VRVCEKMAKYDVVSTLTPQAVMVGSSYGFQYSGFQGRVEFLVNAWAKKCPMGFAYOTRCPDS 1620	
DB	1561 VRVCEKMAKYDVVSTLTPQAVMVGSSYGFQYSGFQGRVEFLVNAWAKKCPMGFAYOTRCPDS 1620	



QY 1621 TWTENDIRVESIYOCCLAPPEARQAIISLTERLYIGGLTNKQNGCYRRCRASGVL 1680  
 Db 1621 TWTENDIRVESIYOCCLAPPEARQAIISLTERLYIGGLTNKQNGCYRRCRASGVL 1680  
 QY 1681 TSCGNLTTCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
 Db 1681 TSCGNLTTCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
 QY 1741 SAPGDPKPEYDLELITSCSSNVSAHDASGRVYVYLTROPTTFLARAANETARHTPEVN 1800  
 Db 1741 SAPGDPKPEYDLELITSCSSNVSAHDASGRVYVYLTROPTTFLARAANETARHTPEVN 1800  
 QY 1801 SWLGNIIWYAPTLWARMILMTHFTSILLQAQLEKALDCQYGCYSTEPDLQIQL 1860  
 Db 1801 SWLGNIIWYAPTLWARMILMTHFTSILLQAQLEKALDCQYGCYSTEPDLQIQL 1860  
 QY 1861 HGLSAPSLHYSPEGEINRVASCLRLKGLVPPRLVRHRSVRARLLSQGGRAAATCGKYL 1920  
 Db 1861 HGLSAPSLHYSPEGEINRVASCLRLKGLVPPRLVRHRSVRARLLSQGGRAAATCGKYL 1920  
 QY 1921 NWAVRTKLTPIPAASQLDLSWFWAGYSGDDIYHLSLRAPRPFWMCLLLSVGVGIY 1980  
 Db 1921 NWAVRTKLTPIPAASQLDLSWFWAGYSGDDIYHLSLRAPRPFWMCLLLSVGVGIY 1980  
 QY 1981 LLEPNR 1985  
 Db 1981 LLEPNR 1985  
 RESULT 12  
 ABG32460  
 ID ABG32460 standard; protein; 3010 AA.  
 XX  
 AC ABG32460;  
 XX  
 DT 15-NOV-2002 (first entry)  
 DE Hepatitis C virus Con 1 isolate polyprotein mutant #9.  
 XX  
 KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 XX internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.  
 OS Hepatitis C virus.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 2199 /note= "Wild-type Ala substituted by Thr"  
 FT  
 XX WO200259321-A2.  
 XX  
 XX 01-AUG-2002.  
 XX  
 XX 16-JAN-2002; 2002WO-BP000526.  
 XX  
 XX 23-JAN-2001; 2001US-0263479P.  
 XX  
 XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.  
 XX  
 XX De Francesco R, Migliaccio G, Paonessa G;  
 XX WPI; 2002-599793/64.  
 XX  
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 PT expression.  
 XX  
 XX Claim 1; Page; 69pp; English.  
 XX  
 XX The invention relates to nucleic acid molecules comprising altered HCV

CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1  
 CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
 CC and NS5B proteins), NS5A mutant of the invention. Note: The present  
 CC sequence is not shown in the specification but was created by the indexer  
 CC using the HCV sequence appearing as ABG32451 and the information in claim  
 CC 1  
 XX  
 SQ Sequence 3010 AA;  
 Query Match 99.9%; Score 10458; DB 5; Length 3010;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1983; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPITAYSOOTRGILGCIITSLTGRDRNOVEGEVQVVTATQSFATCVNGVCMVTVHGA 60  
 Db 1026 LAPITAYSOOTRGILGCIITSLTGRDRNOVEGEVQVVTATQSFATCVNGVCMVTVHGA 1085  
 QY 61 GSKTLAGPKGITQMTYNVDQDLVGMQAPPGARSLTPTCTGSSDLXVLRHADVPVRRR 120  
 Db 1086 GSKTLAGPKGITQMTYNVDQDLVGMQAPPGARSLTPTCTGSSDLXVLRHADVPVRRR 1145  
 QY 121 GDSRGSLSPRPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVEMTTM 180  
 Db 1146 GDSRGSLSPRPVSYLKGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVEMTTM 1205  
 QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTSGKSTKVPAAVAAQGYKVLNPSVAATLGF 240  
 Db 1206 RSPVFTDNSPPAVPQTFQVAHLHAPTSGKSTKVPAAVAAQGYKVLNPSVAATLGF 1265  
 QY 241 AYMSKAHGIDPNIRTCVTRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 300  
 Db 1266 AYMSKAHGIDPNIRTCVTRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 1325  
 QY 301 ILGIGTVLDQAEATAGARLVVLATATPPGSVTVPHNIEEVALSSTGEIIFYGKAIPETI 360  
 Db 1326 ILGIGTVLDQAEATAGARLVVLATATPPGSVTVPHNIEEVALSSTGEIIFYGKAIPETI 1385  
 QY 361 KGRHLIFCHSKKKKCDLAAKLSGLGLNNAVAYYRGLDVSVIPTSGLDVIWATDALMTGFT 420  
 Db 1386 KGRHLIFCHSKKKKCDLAAKLSGLGLNNAVAYYRGLDVSVIPTSGLDVIWATDALMTGFT 1445  
 QY 421 GDFDSVIDCNTCTQTVDFSLDPTTETTTVPQDAVSRQRGRMGYIFRVTPG 480  
 Db 1446 GDFDSVIDCNTCTQTVDFSLDPTTETTTVPQDAVSRQRGRMGYIFRVTPG 1505  
 QY 481 ERPSGMFDSVLCYDAGCAWYELTPAETSVRLEAYLNTPLPCVQDHLFEWESVFTGL 540  
 Db 1506 ERPSGMFDSVLCYDAGCAWYELTPAETSVRLEAYLNTPLPCVQDHLFEWESVFTGL 1565  
 QY 541 THIDAHFLSQTQAGDNFFYLVAQATVCARAQAPPSPSDQWKKLIRLKPFLHPTPL 600  
 Db 1566 THIDAHFLSQTQAGDNFFYLVAQATVCARAQAPPSPSDQWKKLIRLKPFLHPTPL 1625  
 QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVI 660  
 Db 1626 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVI 1685

QY 661 IILSGKPAIPDRVILVREFDEMEBECASHLPYIEQGMQLARQFKQAKIIGLIQTATKQARA 720  
 Db 1686 IILSGKPAIPDRVILVREFDEMEBECASHLPYIEQGMQLARQFKQAKIIGLIQTATKQARA 1745  
 QY 721 AAPVVESKWRTEAPFAWKHMNFISGIQYLAGSLTLCNPAIASLMAFTASITSLPTTQH 780  
 Db 1746 AAPVVESKWRTEAPFAWKHMNFISGIQYLAGSLTLCNPAIASLMAFTASITSLPTTQH 1805  
 QY 781 TILFNILGWWAAQALAPPASAFYAGAGIAGAAVSGIGLGKVLVDILAGYGAGVAGALVA 840  
 Db 1806 TILFNILGWWAAQALAPPASAFYAGAGIAGAAVSGIGLGKVLVDILAGYGAGVAGALVA 1865  
 QY 841 PKWMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGCGAVQMMNRLIAFASR 900  
 Db 1866 PKWMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGCGAVQMMNRLIAFASR 1925  
 QY 901 GNHVSPTHVPESDAAARVTQILSSLTITQLLKRHLHOWINEDCSTPCSGSWLRDVMWIC 960  
 Db 1926 GNHVSPTHVPESDAAARVTQILSSLTITQLLKRHLHOWINEDCSTPCSGSWLRDVMWIC 1985  
 QY 961 TVLTDFTKWLQSKLLPRLPGVFFFCQCGYKGVWRGDMQIMOTPCGAQITGHVKNXGMR 1020  
 Db 1986 TVLTDFTKWLQSKLLPRLPGVFFFCQCGYKGVWRGDMQIMOTPCGAQITGHVKNXGMR 2045  
 QY 1021 IVGPRTCSNTWGTGTPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDEHYVTGM 1080  
 Db 2046 IVGPRTCSNTWGTGTPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDEHYVTGM 2105  
 QY 1081 TTDNVKPCQVPAPEFFTEVDGVLRLHRYAPACKPLLRREVTFLVGLNQVLVGSQLPCEPE 1140  
 Db 2106 TTDNVKPCQVPAPEFFTEVDGVLRLHRYAPACKPLLRREVTFLVGLNQVLVGSQLPCEPE 2165  
 QY 1141 PDVAVLTSMULTDPSHITATKRLARGSPPSIASSASQLSAPSLKATCTTRHDSFAD 1200  
 Db 2166 PDVAVLTSMULTDPSHITATKRLARGSPPSIASSASQLSAPSLKATCTTRHDSFAD 2225  
 QY 1201 LIEANLLWQEMGNTTRVESENKVLILDSFELQAEEDEREVSVAEILRLSRKFPFRAM 1260  
 Db 2226 LIEANLLWQEMGNTTRVESENKVLILDSFELQAEEDEREVSVAEILRLSRKFPFRAM 2285  
 QY 1261 PIWARDYNPPLLESKNDPDPVPPVHGCPLPAKAPPIPPRRKRTVILSESTVSALA 1320  
 Db 2286 PIWARDYNPPLLESKNDPDPVPPVHGCPLPAKAPPIPPRRKRTVILSESTVSALA 2345  
 QY 1321 ELATKTFGSSSAVDSGTATSPDQPSDDGAGSDVESYSSMPLEGEFGDPLDSDGSW 1380  
 Db 2346 ELATKTFGSSSAVDSGTATSPDQPSDDGAGSDVESYSSMPLEGEFGDPLDSDGSW 2405  
 QY 1381 STVSEASESDVCCSMSTYWTGALITPCAEEETKLPIINALSNSILLRHNLVYATTGRSAS 1440  
 Db 2406 STVSEASESDVCCSMSTYWTGALITPCAEEETKLPIINALSNSILLRHNLVYATTGRSAS 2465  
 QY 1441 LRQKVTFFDLQVLDHYRDVLKEMKAKASTVAKLLSVBEACKLTPPHSARSKFYGYAK 1500  
 Db 2466 LRQKVTFFDLQVLDHYRDVLKEMKAKASTVAKLLSVBEACKLTPPHSARSKFYGYAK 2525  
 QY 1501 DVRLNLSKAVNHRSVWVKLLLEDTPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
 Db 2526 DVRLNLSKAVNHRSVWVKLLLEDTPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 2585  
 QY 1561 VRCEKMAVYDVVSTLPOAVNGSSYGFQSPGQRFVFLVNAWKAKCPMGFAYDTRCFDS 1620  
 Db 2586 VRCEKMAVYDVVSTLPOAVNGSSYGFQSPGQRFVFLVNAWKAKCPMGFAYDTRCFDS 2645  
 QY 1621 TVTENDIRVEESIIQCCDLAEPARQAIRSLTERLYIGPLTNSKGQNCVYRCRASGVL 1680  
 Db 2646 TVTENDIRVEESIIQCCDLAEPARQAIRSLTERLYIGPLTNSKGQNCVYRCRASGVL 2705  
 QY 1681 TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLAFTEAMTRY 1740  
 Db 2706 TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLAFTEAMTRY 2765

QY 1741 SAPGDPKPKPYDLELITSCSSNVSVADHAGSKRVYVLTROPTTPLAANAETARHTPVN 1800  
 Db 2766 SAPGDPKPKPYDLELITSCSSNVSVADHAGSKRVYVLTROPTTPLAANAETARHTPVN 2825  
 QY 1801 SWLGNIMYAPTILWARMILMTHFFSILLAQSEKALDCQIYGACYSIEPLDLPOLIQL 1860  
 Db 2826 SWLGNIMYAPTILWARMILMTHFFSILLAQSEKALDCQIYGACYSIEPLDLPOLIQL 2885  
 QY 1861 HGLSAFSLHSYSPGEINRVASCLRKLGVPLRVHRHARSVRARLLSQGGAATCGKYL 1920  
 Db 2886 HGLSAFSLHSYSPGEINRVASCLRKLGVPLRVHRHARSVRARLLSQGGAATCGKYL 2945  
 QY 1921 NNAVRTKLKLPPIPAASQDLSSWVFAGYSGDIYHLSLRARPRFWMCLLLSVGVGIY 1980  
 Db 2946 NNAVRTKLKLPPIPAASQDLSSWVFAGYSGDIYHLSLRARPRFWMCLLLSVGVGIY 3005  
 QY 1981 LLPNR 1985  
 Db 3006 LLPNR 3010

RESULT 13  
 ABG32454  
 ID ABG32454 standard; protein; 3010 AA.  
 XX  
 AC ABG32454;  
 XX  
 DT 15-NOV-2002 (first entry)  
 XX  
 DE Hepatitis C virus Con 1 isolate polyprotein mutant #3.  
 XX  
 KW HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutain;  
 KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
 KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.  
 XX  
 OS Hepatitis C virus.  
 OS Synthetic.  
 XX  
 XX Key Location/Qualifiers  
 FH Misc-difference 1347 /note= "Wild-type Ala substituted by Thr"  
 FT  
 FT  
 XX  
 XX WO200259321-A2.  
 XX  
 XX 01-AUG-2002.  
 XX  
 XX 16-JAN-2002; 2002WO-BF000526.  
 XX  
 XX 23-JAN-2001; 2001US-0263479P.  
 XX  
 XX (RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.  
 XX  
 XX De Francesco R, Migliaccio G, Paonessa G;  
 XX  
 XX WPI; 2002-599793/64.  
 XX  
 XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
 PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
 PT ribosome entry site (IRES) region, useful in studying HCV replication and  
 PT expression.  
 XX  
 XX Claim 1; Page: 69pp; English.  
 XX  
 XX The invention relates to nucleic acid molecules comprising altered HCV  
 CC NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
 CC internal ribosome entry site (IRES) region coding for one or more NS3,  
 CC NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
 CC are detailed in the specification. Also included are (1) an expression  
 CC vector comprising a nucleotide sequence coding for the altered nucleic  
 CC acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
 CC recombinant cell human hepatoma cell comprising the altered nucleic acids  
 CC ; (3) a recombinant cell produced by introducing into a human hepatoma  
 CC cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)

CC replicon enhanced cell or which containing a functional HCV replicon; (5)  
 CC an HCV replicon enhanced cells made in the method; and (6) measuring the  
 CC ability of a compound to affect HCV activity. The HCV replicons and HCV  
 CC replicon enhanced cells are useful in studying HCV replication and  
 CC expression, and HCV and host cell interactions, producing HCV RNA and  
 CC proteins, and providing a system for measuring the ability of a compound  
 CC to modulate one or more HCV activities e.g. to discover drugs which may  
 CC treat HCV mediated diseases such as liver failure, cirrhosis and  
 CC hepatocellular carcinoma. The present sequence is the HCV replicon Con 1  
 CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
 CC and NS5B proteins), NS3 mutant of the invention. Note: The present  
 CC sequence is not shown in the specification but was created by the indexer  
 CC using the HCV sequence appearing as ABG32451 and the information in Claim  
 CC 1  
 XX  
 SQ Sequence 3010 AA;

Query Match 99.9%; Score 10458; DB 5; Length 3010;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1983; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPITAYSOQTRGLLGLIISLTGRDRNQVEGEVQVSTATQSFATCVNGVCVTVYHGA 60  
 Db :  
 1026 LAPITAYSOQTRGLLGLIISLTGRDRNQVEGEVQVSTATQSFATCVNGVCVTVYHGA 1085

QY 61 GSKTLAGPKGPIITOMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLYLVTRHADVIPVRRR 120  
 Db GSKTLAGPKGPIITOMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLYLVTRHADVIPVRRR 1145

QY 121 GDSRGSLLSPRVSYLYKSGSGGPLLCPSPGSHAVGIFRAAVCTRGVAXAVDFVPVSMETTM 180  
 Db GDSRGSLLSPRVSYLYKSGSGGPLLCPSPGSHAVGIFRAAVCTRGVAXAVDFVPVSMETTM 1205

QY 181 RSPVFTNSPPAVPQTFQVAHLHAPTGSQSKTKVPAAYAAQYKVLVLPNSVAATLFGF 240  
 Db RSPVFTNSPPAVPQTFQVAHLHAPTGSQSKTKVPAAYAAQYKVLVLPNSVAATLFGF 1265

QY 241 AYMSKAGHDIPNRTGVRTITTCAPITYSYVYKFLADGGCGGAYDIIICDECHSDSTT 300  
 Db AYMSKAGHDIPNRTGVRTITTCAPITYSYVYKFLADGGCGGAYDIIICDECHSDSTT 1325

QY 301 ILGIGTVLDOAETAGARLVVLTATPPGSVTVPHNPNEEVALSSTGEIPIFYGKAIPLETI 360  
 Db ILGIGTVLDOAETAGARLVVLTATPPGSVTVPHNPNEEVALSSTGEIPIFYGKAIPLETI 1385

QY 361 KGRGHLIFCHSKKKDELAALKSLGLNAVAYYRGDLVSVIPSGDVIVVATDALMTGFT 420  
 Db KGRGHLIFCHSKKKDELAALKSLGLNAVAYYRGDLVSVIPSGDVIVVATDALMTGFT 1445

QY 421 GPDFSVIDCMTCTVQTVDFSLDPTFTIETTVPQDAVSRSGRGTGRGMGIYRFVTPG 480  
 Db GPDFSVIDCMTCTVQTVDFSLDPTFTIETTVPQDAVSRSGRGTGRGMGIYRFVTPG 1505

QY 481 ERPSGMFDSVLCYCEYDAGCAYELTTPAETSVRLRAYLNTPLGIPVQDHLFEWESVFTGL 540  
 Db ERPSGMFDSVLCYCEYDAGCAYELTTPAETSVRLRAYLNTPLGIPVQDHLFEWESVFTGL 1565

QY 541 THIDAHFLSQTQAGNFPYLVAQYATVCARAAQPPSDQMWKCLIRLKTPLHGTPTLL 600  
 Db THIDAHFLSQTQAGNFPYLVAQYATVCARAAQPPSDQMWKCLIRLKTPLHGTPTLL 1625

QY 601 YRLGAVQNEVTTTHPTTKYIMACMSADLEVTSTWLVGVLAALAAAYCLTTGSSVIVGR 660  
 Db YRLGAVQNEVTTTHPTTKYIMACMSADLEVTSTWLVGVLAALAAAYCLTTGSSVIVGR 1685

QY 661 IILSGKPAIIPDREVLRYREFDEMEECASHLPYIEQWQLAEQFKQKAIQLQTATKQAEA 720  
 Db IILSGKPAIIPDREVLRYREFDEMEECASHLPYIEQWQLAEQFKQKAIQLQTATKQAEA 1745

QY 721 AAPVVESKRWRTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASMAFTASITSLTTOH 780  
 Db AAPVVESKRWRTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASMAFTASITSLTTOH 1805

QY 781 TLLFNILGGWVAQAAPPASAASFVAGIAGAAVGSIGLKVLDVILAGYGAGVAGALVA 840  
 Db TLLFNILGGWVAQAAPPASAASFVAGIAGAAVGSIGLKVLDVILAGYGAGVAGALVA 1865

QY 841 FKVMSGEMPTSEDLVNLIPAILSPGALVGVVCAAILRRHVGPGEVQVMMNRLIAFASR 900  
 Db FKVMSGEMPTSEDLVNLIPAILSPGALVGVVCAAILRRHVGPGEVQVMMNRLIAFASR 1925

QY 901 GNVSPHYVPESDAAARVTOILSSLTITQLKRLHOMINEDCSTPCSGSMLRDVMDIC 960  
 Db GNVSPHYVPESDAAARVTOILSSLTITQLKRLHOMINEDCSTPCSGSMLRDVMDIC 1985

QY 961 TVLTFDKTWLQSKLLPRLPGVPFFSCQYKGVWRGDMOTTCPCGACIITHGVKNGSMR 1020  
 Db TVLTFDKTWLQSKLLPRLPGVPFFSCQYKGVWRGDMOTTCPCGACIITHGVKNGSMR 2045

QY 1021 IVGPRTCSNTHWGTFPINAITYTGTCTSPAPNYGRALWRVAABEYVEVTRVGDPHYVTGM 1080  
 Db IVGPRTCSNTHWGTFPINAITYTGTCTSPAPNYGRALWRVAABEYVEVTRVGDPHYVTGM 2105

QY 1081 TTDNVKCPQVPAPEFFTEVDGVELHRYAPACKLLREEVTFVLGMLNOYLVGSLPCEPE 1140  
 Db TTDNVKCPQVPAPEFFTEVDGVELHRYAPACKLLREEVTFVLGMLNOYLVGSLPCEPE 2165

QY 1141 PDVAVLTSMLTDPESHITAEAKRLARSPSPSLASSASQSLSPSLKATCTTRHDSPPAD 1200  
 Db PDVAVLTSMLTDPESHITAEAKRLARSPSPSLASSASQSLSPSLKATCTTRHDSPPAD 2225

QY 1201 LIEANLLWROEMGNIITRVSENKVILDSFEPLQAEDEREVSVPABILRRSRKFFRAM 1260  
 Db LIEANLLWROEMGNIITRVSENKVILDSFEPLQAEDEREVSVPABILRRSRKFFRAM 2285

QY 1261 PIWARPDPYNPLLESWKDPDYVPVHVHGCPLPAKAPPIPPBRKRVTVLSESTVSALA 1320  
 Db PIWARPDPYNPLLESWKDPDYVPVHVHGCPLPAKAPPIPPBRKRVTVLSESTVSALA 2345

QY 1321 ELATKTFGSESSAIVSGTATASPDQSPDDGAGSDVESYSSMPPLGEPEGDPDLSGWS 1380  
 Db ELATKTFGSESSAIVSGTATASPDQSPDDGAGSDVESYSSMPPLGEPEGDPDLSGWS 2405

QY 1381 STVSEESADVDVCCSYTWTGALITPCAABETKLPINALNSLLRHNNIVYATTSRAS 1440  
 Db STVSEESADVDVCCSYTWTGALITPCAABETKLPINALNSLLRHNNIVYATTSRAS 2465

QY 1441 LRQKVTFRQLQVLDHRYDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYAK 1500  
 Db LRQKVTFRQLQVLDHRYDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYAK 2525

QY 1501 DVNLSKSAVNHIRSVWKOLLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
 Db DVNLSKSAVNHIRSVWKOLLEDETETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 2585

QY 1561 VRVCEKALYDVVSTLPQAWMGSSYGFQSPQORVEFLVNAWKAKCPMGFAVDTTCFDS 1620  
 Db VRVCEKALYDVVSTLPQAWMGSSYGFQSPQORVEFLVNAWKAKCPMGFAVDTTCFDS 2645

QY 1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGGLTNSKGNCCYVRCRAGSVLT 1680  
 Db TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGGLTNSKGNCCYVRCRAGSVLT 2705

QY 1681 TSCGNLTCLYKAAACRAAKLQDCTMLVCGDDLTVVICESAGTOEDEASLRAFTAMTRY 1740  
 Db TSCGNLTCLYKAAACRAAKLQDCTMLVCGDDLTVVICESAGTOEDEASLRAFTAMTRY 2765

QY 1741 SAPPDGPPEYDLEILITSCSSNVSAHDASGRVYLYLTRDPTTPLARAAMTARHTPN 1800  
 Db SAPPDGPPEYDLEILITSCSSNVSAHDASGRVYLYLTRDPTTPLARAAMTARHTPN 2825

QY 1801 SWLGNIMYAPTLWARMILMTHFFSILLAOEQLKALDCQIYGACYSLEPDLPOIQL 1860  
 Db SWLGNIMYAPTLWARMILMTHFFSILLAOEQLKALDCQIYGACYSLEPDLPOIQL 2885

QY 1861 HGLSAFSLHSYSPGEBINRVAACLKXGLGVPLRVWRHRARSVRARLLSOGGRAATCGKLYF 1920

Db 2886 HGLSAFSLHSYDPEINRVAACLRKLGVPPIURVWRHRSVRARLLSOGGRRATCGKLYF 2945  
QY 1921 NWAVRTKLLPIPAASOLDLSSWFVAGYSGGDIYHSLSRARPRFWMCLLLLSVGVGIY 1980  
Db 2946 NWAVRTKLLPIPAASOLDLSSWFVAGYSGGDIYHSLSRARPRFWMCLLLLSVGVGIY 3005

QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

RESULT 14  
ABG32461  
ID ABG32461 standard; protein; 3010 AA.

XX AC ABG32461;

DT 15-NOV-2002 (first entry)

XX Hepatitis C virus Con 1 isolate polyprotein mutant #10.

XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; muten; HCV;  
XX hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
XX internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.

OS Hepatitis C virus.  
OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 2204 /note= "Wild-type Ser substituted by Arg"

XX WO200259321-A2.

XX 01-AUG-2002.

XX 16-JAN-2002; 2002WO-EP000526.

XX 23-JAN-2001; 2001US-0263479P.

XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.

XX De Francesco R, Migliaccio G, Paonessa G;

XX MPI; 2002-599793/64.

XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
XX ribosome entry site (IRES) region, useful in studying HCV replication and  
XX expression.

XX Claim 1; Page; 69pp; English.

XX The invention relates to nucleic acid molecules comprising altered HCV  
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
XX internal ribosome entry site (IRES) region coding for one or more NS3,  
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
XX are detailed in the specification. Also included are (1) an expression  
XX vector comprising a nucleotide sequence coding for the altered nucleic  
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
XX recombinant cell human hepatoma cell comprising the altered nucleic acids  
XX ; (3) a recombinant cell produced by introducing into a human hepatoma  
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
XX replicon enhanced cell or which containing a functional HCV replicon; (5)  
XX an HCV replicon enhanced cells made in the method; and (6) measuring the  
XX ability of a compound to affect HCV activity. The HCV replicons and HCV  
XX replicon enhanced cells are useful in studying HCV replication and  
XX expression, and HCV and host cell interactions, producing HCV RNA and  
XX proteins, and providing a system for measuring the ability of a compound  
XX to modulate one or more HCV activities e.g. to discover drugs which may  
XX treat HCV mediated diseases such as liver failure, cirrhosis and  
XX hepatocellular carcinoma. The present sequence is the HCV replicon Con 1

CC polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
CC and NS5B proteins). NS5A mutant of the invention. Note: The present  
CC sequence is not shown in the specification but was created by the indexer  
CC using the HCV sequence appearing as ABG32451 and the information in claim  
CC 1  
XX SQ Sequence 3010 AA;

Query Match 99.9%; Score 10457; DB 5; Length 3010;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1983; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MAPITAYSQQTGRGLGCIITSLTGRDRNQVGEVQVWSTATQSFLATCVNGVCTWTVYHGA 60  
Db 1026 LAPITAYSQQTGRGLGCIITSLTGRDRNQVGEVQVWSTATQSFLATCVNGVCTWTVYHGA 1085  
QY 61 GSKTLAGPKGPITQMYTNVDQDLVQWQAPPGARSILTPCTCGSSDLVLYVTRHADVLPVRRR 120  
Db 1086 GSKTLAGPKGPITQMYTNVDQDLVQWQAPPGARSILTPCTCGSSDLVLYVTRHADVLPVRRR 1145  
QY 121 GDSRGSLLSPRPVSYLKSGSGGPLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 1146 GDSRGSLLSPRPVSYLKSGSGGPLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205  
QY 181 RSPVFTDNSSPPAVPQTQVQVHLHAPTSGSKSTKVPAAVAAQYKVLVLPNSVAATLGF 240  
Db 1206 RSPVFTDNSSPPAVPQTQVQVHLHAPTSGSKSTKVPAAVAAQYKVLVLPNSVAATLGF 1265  
QY 241 AYMSKAHGIDENIRTCVTRITTTGAPITYSTYGFADGCGSGAYDIIICDECHSDSTT 300  
Db 1266 AYMSKAHGIDENIRTCVTRITTTGAPITYSTYGFADGCGSGAYDIIICDECHSDSTT 1325  
QY 301 ILGIGTVLDQARTAGARLVLAATATPPGSVTVPHNIEEVALSSTGEIPFYGKAIPETI 360  
Db 1326 ILGIGTVLDQARTAGARLVLAATATPPGSVTVPHNIEEVALSSTGEIPFYGKAIPETI 1385  
QY 361 KGRHLIFCHSKKKCDELAALKSGLNNAVAYRGDLVSVITPSGDIIVADALMTGT 420  
Db 1386 KGRHLIFCHSKKKCDELAALKSGLNNAVAYRGDLVSVITPSGDIIVADALMTGT 1445  
QY 421 GDFDSVIDCNTCTQTVDFSLDPTFTIETTVQDAVSRSQRRGTRGRMGIRYRVTPG 480  
Db 1446 GDFDSVIDCNTCTQTVDFSLDPTFTIETTVQDAVSRSQRRGTRGRMGIRYRVTPG 1505  
QY 481 ERPSGMFDSVLCCEYDAGCAWYELTPAETSVELRAYLNTPLGPVCDHLFPWESVFTGL 540  
Db 1506 ERPSGMFDSVLCCEYDAGCAWYELTPAETSVELRAYLNTPLGPVCDHLFPWESVFTGL 1565  
QY 541 THIDAHFLSQTQAGDNFPYLVAYQATVCARAQAPPPSWDQWKKLIRLKPDLHGPTPL 600  
Db 1566 THIDAHFLSQTQAGDNFPYLVAYQATVCARAQAPPPSWDQWKKLIRLKPDLHGPTPL 1625  
QY 601 YRLGAVONEVTTTHPIITKYMCMASADLEVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 660  
Db 1626 YRLGAVONEVTTTHPIITKYMCMASADLEVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 1685  
QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLAEQFKQKGAIGLLQATQAE 720  
Db 1686 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLAEQFKQKGAIGLLQATQAE 1745  
QY 721 AAPVSKMTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSELTQH 780  
Db 1746 AAPVSKMTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSELTQH 1805  
QY 781 TLLFNILGGVWAQALAPPSSAASAFVAGIAGAAGVSTGLGKLVLDVILAGYGAGVAGALVA 840  
Db 1806 TLLFNILGGVWAQALAPPSSAASAFVAGIAGAAGVSTGLGKLVLDVILAGYGAGVAGALVA 1865  
QY 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 900  
Db 1866 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 1925  
QY 901 GNHVSPTHYVPESDAAARVQTILSSLLTITQLLKHQLWINEDECSTPCSGSWLRDWDWIC 960

Db 1926 GNVSTHVVPESDAARVTOILSSITITQLLKRLHQWINEDESTCSCSWLRDWDWIC 1985  
QY 961 TVLTDFTKTLQSKLLPRLPGVFFSCQGYKYGVRGDMQITTCPCGAQITGHVKNQSMR 1020  
Db 1986 TVLTDFTKTLQSKLLPRLPGVFFSCQGYKYGVRGDMQITTCPCGAQITGHVKNQSMR 2045  
QY 1021 INVPRCSTWHTGTFPINAITYTTCPTSPAPNYSRALNRVAABEYVEVTRVGDHYVTGM 1080  
Db 2046 INVPRCSTWHTGTFPINAITYTTCPTSPAPNYSRALNRVAABEYVEVTRVGDHYVTGM 2105  
QY 1081 TTDNVKCPQVPAPEPTEVDGVRHRYAPACKPLLEEVTVLVGLNQYLVSQSLCEPE 1140  
Db 2106 TTDNVKCPQVPAPEPTEVDGVRHRYAPACKPLLEEVTVLVGLNQYLVSQSLCEPE 2165  
QY 1141 PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLIASSASQSLAPSILKATCTTRHDSPAD 1200  
Db 2166 PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLIASSASQSLAPSILKATCTTRHDSPAD 2225  
QY 1201 LIEANLLWQEMGNITRVESKNVILDSFELQAEEDEREVSVAEILRRSRKFPFRAM 1260  
Db 2226 LIEANLLWQEMGNITRVESKNVILDSFELQAEEDEREVSVAEILRRSRKFPFRAM 2285  
QY 1261 PIWAPDYNPPLLESWKDPDYVPPVHVGCPPLPAKAPPIPPRRKRTVVLSESTVSALA 1320  
Db 2286 PIWAPDYNPPLLESWKDPDYVPPVHVGCPPLPAKAPPIPPRRKRTVVLSESTVSALA 2345  
QY 1321 ELATKTFGSSSASVDSGTATASPDQSDGDAGSDVESYSSMPPLGEPGPDLSGWS 1380  
Db 2346 ELATKTFGSSSASVDSGTATASPDQSDGDAGSDVESYSSMPPLGEPGPDLSGWS 2405  
QY 1381 STVSEASDVVCCSMYSYTWGALITPCAETKLPINALSNLSLRLHNLVYATTSRSAS 1440  
Db 2406 STVSEASDVVCCSMYSYTWGALITPCAETKLPINALSNLSLRLHNLVYATTSRSAS 2465  
QY 1441 LRQKVTTFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFGYGA 1500  
Db 2466 LRQKVTTFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFGYGA 2525  
QY 1501 DVYRNLSSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKRPALIVFPD 1560  
Db 2526 DVYRNLSSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGGKRPALIVFPD 2585  
QY 1561 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQVBEFLVNAWKAKKCPMGFYDTRCFDS 1620  
Db 2586 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQVBEFLVNAWKAKKCPMGFYDTRCFDS 2645  
QY 1621 TVTENDIRVEESTIYQCCDLAPARQAIRSLTERLYIGGLPLTNSKGONCCGYRRCRAGSVLT 1680  
Db 2646 TVTENDIRVEESTIYQCCDLAPARQAIRSLTERLYIGGLPLTNSKGONCCGYRRCRAGSVLT 2705  
QY 1681 TSCGNLTLCYLAACAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740  
Db 2706 TSCGNLTLCYLAACAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 2765  
QY 1741 SAPPDGPPEVDELEITSCSNVSVAHDASGRVYVLTTRDPTPLARAAMTARHTPVN 1800  
Db 2766 SAPPDGPPEVDELEITSCSNVSVAHDASGRVYVLTTRDPTPLARAAMTARHTPVN 2825  
QY 1801 SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIBRL 1860  
Db 2826 SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIBRL 2885  
QY 1861 HGLSAPLSHSYSGEINRVASCLRLGVPLPLVRHRRARSVRARLLSQGGRAATCGKYL 1920  
Db 2886 HGLSAPLSHSYSGEINRVASCLRLGVPLPLVRHRRARSVRARLLSQGGRAATCGKYL 2945  
QY 1921 NNAVTKLTLPIPAASQDLSSWFVAGYSGGDIYHLSRARPFRMFWCMLLLSVGVGIY 1980  
Db 2946 NNAVTKLTLPIPAASQDLSSWFVAGYSGGDIYHLSRARPFRMFWCMLLLSVGVGIY 3005  
QY 1981 LILPNR 1985  
| | | |

Db 3006 LILPNR 3010  
RESULT 15  
ABG32458  
ID ABG32458 standard; protein; 3010 AA.  
XX ABG32458;  
XX AC  
XX DT 15-NOV-2002 (first entry)  
XX Hepatitis C virus Con 1 isolate polyprotein mutant #7.  
XX HCV; Con 1; adaptive mutation; liver failure; cirrhosis; mutant; mutain;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRES; NS5A; HCV replication; polyprotein.  
XX Hepatitis C virus.  
OS Synthetic.  
XX  
XX Key Location/Qualifiers  
FH Misc-difference 2197  
FT /note= "Wild-type Phe substituted by Ser"  
XX WO200259321-A2.  
XX 01-AUG-2002.  
XX 16-JAN-2002; 2002WO-BP000526.  
XX 23-JAN-2001; 2001US-0263479P.  
XX (RICE-) IST RICERCH BIOL MOLECOLARE ANGELETTI.  
XX De Francesco R, Migliaccio G, Paonessa G;  
XX WPI; 2002-599793/64.  
PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX Claim 1; Page; 69pp; English.  
XX The invention relates to nucleic acid molecules comprising altered HCV  
NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
internal ribosome entry site (IRES) region coding for one or more NS3,  
NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
are detailed in the specification. Also included are (1) an expression  
vector comprising a nucleotide sequence coding for the altered nucleic  
acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
recombinant cell human hepatoma cell comprising the altered nucleic acids  
; (3) a recombinant cell produced by introducing into a human hepatoma  
cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
replicon enhanced cell or which containing a functional HCV replicon; (5)  
an HCV replicon enhanced cells made in the method; and (6) measuring the  
ability of a compound to affect HCV activity. The HCV replicons and HCV  
replicon enhanced cells are useful in studying HCV replication and  
expression, and HCV and host cell interactions, producing HCV RNA and  
proteins, and providing a system for measuring the ability of a compound  
to modulate one or more HCV activities e.g. to discover drugs which may  
treat HCV mediated diseases such as liver failure, cirrhosis and  
hepatocellular carcinoma. The present sequence is the HCV replicon Con 1  
polyprotein (comprising the Core, E1, E2, P7, NS2, NS3, NS4A, NS4B, NS5A  
and NS5B proteins), NS5A mutant of the invention. Note: The present  
sequence is not shown in the specification but was created by the indexer  
using the HCV sequence appearing as ABG32451 and the information in claim  
1  
SQ Sequence 3010 AA;

Query Match 99.9%; Score 10456; DB 5; Length 3010;





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:20:57 ; Search time 33 Seconds  
(without alignments)  
3989.128 Million cell updates/sec

Title: US-09-576-989-3  
Perfect score: 10465  
Sequence: 1 MAPITAYSQTRGLGCIIT.....FMWCLLLLSVGVGVYLLPNR 1985

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10465	100.0	1985	4	US-09-539-601-9
2	10465	100.0	1985	4	US-09-539-601-12
3	10462	100.0	2201	4	US-09-539-601-6
4	10462	100.0	2201	4	US-09-539-601-15
5	10462	100.0	3010	4	US-09-539-601-3
6	10442	99.8	2201	4	US-10-029-907-3
7	10432	99.7	1985	4	US-09-539-601-27
8	10432	99.7	3010	4	US-09-539-601-24
9	10425	99.6	1985	4	US-09-539-601-18
10	10422	99.6	3010	4	US-09-539-601-21
11	10412	99.5	1985	4	US-09-539-601-30
12	10409	99.5	3010	4	US-09-539-601-33
13	10170	97.2	2201	3	US-08-952-981A-2
14	10170	97.2	2621	2	US-08-324-977-36
15	10170	97.2	2621	2	US-08-384-616-36
16	10170	97.2	2621	3	US-09-315-850-36
17	10170	97.2	2621	3	US-09-315-850-36
18	10170	97.2	3010	1	US-08-324-977-2
19	10170	97.2	3010	1	US-08-324-977-14
20	10170	97.2	3010	2	US-08-384-616-2
21	10170	97.2	3010	2	US-08-384-616-14
22	10170	97.2	3010	2	US-08-304-686A-2
23	10170	97.2	3010	2	US-08-304-686A-14
24	10170	97.2	3010	3	US-09-315-850-2
25	10170	97.2	3010	3	US-09-315-850-14
26	10165	97.1	2620	1	US-08-324-977-32
27	10165	97.1	2620	2	US-08-384-616-32

28	10165	97.1	2620	2	US-08-904-686A-32	Sequence 32, Appl
29	10165	97.1	2620	3	US-09-315-850-32	Sequence 32, Appl
30	10106	96.6	3010	3	US-09-014-416-3	Sequence 3, Appl
31	9430.5	90.1	3012	3	US-08-811-566-2	Sequence 2, Appl
32	9430.5	90.1	3012	3	US-09-034-756-2	Sequence 2, Appl
33	9426.5	90.1	3011	1	US-08-453-552-2	Sequence 2, Appl
34	9426.5	90.1	3011	2	US-08-710-637-2	Sequence 2, Appl
35	9426.5	90.1	3011	3	US-08-850-328-1	Sequence 1, Appl
36	9426.5	90.1	3011	5	PCT-US93-00907-2	Sequence 1, Appl
37	9418.5	90.0	3011	3	US-09-014-416-1	Sequence 1, Appl
38	9418.5	90.0	3011	3	US-09-014-416-5	Sequence 5, Appl
39	9417.5	90.0	3011	4	US-09-952-572-9	Sequence 9, Appl
40	9417.5	90.0	3011	1	US-08-440-103-36	Sequence 36, Appl
41	9417.5	90.0	3011	1	US-08-440-542-36	Sequence 36, Appl
42	9417.5	90.0	3011	1	US-07-910-760-10	Sequence 10, Appl
43	9417.5	90.0	3011	1	US-08-440-519-10	Sequence 10, Appl
44	9417.5	90.0	3011	1	US-08-231-368-36	Sequence 36, Appl
45	9417.5	90.0	3011	1	US-08-440-210-36	Sequence 36, Appl

ALIGNMENTS

RESULT 1  
US-09-539-601-9  
; Sequence 9, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 9  
; LENGTH: 1985  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-9

Query Match	100.0%;	Score 10465;	DB 4;	Length 1985;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1985;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MAPITAYSQTRGLGCIITSLTGRDRNQVEGEVQVWVSTATQSEFLATCVNGVCWTVYHGA	60	
Db	1	MAPITAYSQTRGLGCIITSLTGRDRNQVEGEVQVWVSTATQSEFLATCVNGVCWTVYHGA	60	
QY	61	GSKTLAGPKGPITQMTYNNVDQLVGMQAPPGARSILTPCTCGSSDLLVTRHADVIPVRRR	120	
Db	61	GSKTLAGPKGPITQMTYNNVDQLVGMQAPPGARSILTPCTCGSSDLLVTRHADVIPVRRR	120	
QY	121	GSRSGLSPRVSVYKSGSGPLICPSGHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM	180	
Db	121	GSRSGLSPRVSVYKSGSGPLICPSGHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM	180	
QY	181	RSPVFTDNSSPPAVPQTFQVHLHAPTGSKSTKVPAAVAAQGYKVLNPNVAATLGF	240	
Db	181	RSPVFTDNSSPPAVPQTFQVHLHAPTGSKSTKVPAAVAAQGYKVLNPNVAATLGF	240	
QY	241	AYMSXAHGIDPNIRTVGRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT	300	
Db	241	AYMSXAHGIDPNIRTVGRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT	300	
QY	301	ILGIGTVLDQATAGARLVWLATATPPGSTVTPHPNIEVALSSGTGEIPFYKAIPIETI	360	
Db	301	ILGIGTVLDQATAGARLVWLATATPPGSTVTPHPNIEVALSSGTGEIPFYKAIPIETI	360	
QY	361	KGGRHLIFCHSKKCKDELAALKSLGILNAVAYRGIDUSVIETSGDIVVATDALMTGFT	420	



Db 361 KGRHILFCHSKKKDELAALKSLGLGNAYVYRGLDVSVIPTSGDVIWATDALMTGFT 420  
Qy 421 GDFSVDCNTCVCTVDFPSLDPTTETTTVPQDAVRSORRGHGTGRMGIIYFVTPG 480  
Db 421 GDFSVDCNTCVCTVDFPSLDPTTETTTVPQDAVRSORRGHGTGRMGIIYFVTPG 480  
Qy 481 ERPSGMFSSVLCYCYDAGCAWYELTPAETSVRIRAYLNTPLGVPQCQDHLFEWESVFTGL 540  
Db 481 ERPSGMFSSVLCYCYDAGCAWYELTPAETSVRIRAYLNTPLGVPQCQDHLFEWESVFTGL 540  
Qy 541 THIDAHFTSOTQAGDNFPYLVAYQATVCARAQAPPSWDQWKKLIRLKPTLHGPTPLL 600  
Db 541 THIDAHFTSOTQAGDNFPYLVAYQATVCARAQAPPSWDQWKKLIRLKPTLHGPTPLL 600  
Qy 601 YRLGAVQNEVTTTHPIIKYINACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTSVVIVGR 660  
Db 601 YRLGAVQNEVTTTHPIIKYINACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTSVVIVGR 660  
Qy 661 IILSGKPAIIPDREVLVREFDEMBCASHLPYIEQGMQLAEQFKQKAILGLLQTATKQAEA 720  
Db 661 IILSGKPAIIPDREVLVREFDEMBCASHLPYIEQGMQLAEQFKQKAILGLLQTATKQAEA 720  
Qy 721 AAPVVESKWRITLAEAFWAKHMNFISGIQYLAGSLTLPGNPAIASLMAFTASITSPLTQH 780  
Db 721 AAPVVESKWRITLAEAFWAKHMNFISGIQYLAGSLTLPGNPAIASLMAFTASITSPLTQH 780  
Qy 781 TLLFNILGWWAAQALAPPSAFAFGAGIAGAAVGSIGLGLVLDILAGYGAGVAGALVA 840  
Db 781 TLLFNILGWWAAQALAPPSAFAFGAGIAGAAVGSIGLGLVLDILAGYGAGVAGALVA 840  
Qy 841 FKVMGEMPESTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGEVQVWNNRLJAFASR 900  
Db 841 FKVMGEMPESTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGEVQVWNNRLJAFASR 900  
Qy 901 GNVHSPHYVPESDAARVTOILSLTITQLLKRLHQWINEDECSTPCSGSMLRDVMDWIC 960  
Db 901 GNVHSPHYVPESDAARVTOILSLTITQLLKRLHQWINEDECSTPCSGSMLRDVMDWIC 960  
Qy 961 TVLTDFTKWLQSKLLPRLPVGPFFPFCORGKGVWKGDMQTTCPGQAQITGHVKNKSGMR 1020  
Db 961 TVLTDFTKWLQSKLLPRLPVGPFFPFCORGKGVWKGDMQTTCPGQAQITGHVKNKSGMR 1020  
Qy 1021 IVGPRTCSNTWHGTFPPINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDGFHYVTGM 1080  
Db 1021 IVGPRTCSNTWHGTFPPINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDGFHYVTGM 1080  
Qy 1081 TTDNVKCPQVPAPEFFTEVDGVRILHRVAPACKPILREEVFLVGLNQYLVGSQLPCBPPE 1140  
Db 1081 TTDNVKCPQVPAPEFFTEVDGVRILHRVAPACKPILREEVFLVGLNQYLVGSQLPCBPPE 1140  
Qy 1141 PDVAVLTSMITDPGSHITAEAKRRLARGSPPSLASSASQLSAPSLKATCTTRHDSPAD 1200  
Db 1141 PDVAVLTSMITDPGSHITAEAKRRLARGSPPSLASSASQLSAPSLKATCTTRHDSPAD 1200  
Qy 1201 LIEANLLWRQEMGNITRVESENKVVILDSFEPLQAEEDEREVSVPAILRRSRKFFPRAM 1260  
Db 1201 LIEANLLWRQEMGNITRVESENKVVILDSFEPLQAEEDEREVSVPAILRRSRKFFPRAM 1260  
Qy 1261 PIWARPNDYNPLESKWDPDVPVPHVHCPIPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 1261 PIWARPNDYNPLESKWDPDVPVPHVHCPIPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Qy 1321 ELATKTFGSSSSAVDSGTATASDPQSDGSDVESYSMPLEBGPDPDLSGWSW 1380  
Db 1321 ELATKTFGSSSSAVDSGTATASDPQSDGSDVESYSMPLEBGPDPDLSGWSW 1380  
Qy 1381 STVSEASEDVVCCSMSTWTGALITPCAAEBTKLPINALNSLRHNLVYATTSRSAS 1440  
Db 1381 STVSEASEDVVCCSMSTWTGALITPCAAEBTKLPINALNSLRHNLVYATTSRSAS 1440  
Qy 1441 LROKVVTFDRLOVLDHDDYRDLVKEMKAKASTVKALLSVEEBACKLTPPHSARSKEGYGAK 1500  
Db 1441 LROKVVTFDRLOVLDHDDYRDLVKEMKAKASTVKALLSVEEBACKLTPPHSARSKEGYGAK 1500

Qy 1501 DVRLSSKAVNHIRSVWKDLLEDTEPTDITIMAKNEVFCVQPEKGGKRPARLIIVFPDLG 1560  
Db 1501 DVRLSSKAVNHIRSVWKDLLEDTEPTDITIMAKNEVFCVQPEKGGKRPARLIIVFPDLG 1560  
Qy 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPQQRVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
Db 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPQQRVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
Qy 1621 TVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQNGCYRRCRASGVLT 1680  
Db 1621 TVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQNGCYRRCRASGVLT 1680  
Qy 1681 TSCGNLTLCYLKAAAACRAAKLQDCTMLVCGDDLVIWICESAGTQDEASLRAFTAMTRY 1740  
Db 1681 TSCGNLTLCYLKAAAACRAAKLQDCTMLVCGDDLVIWICESAGTQDEASLRAFTAMTRY 1740  
Qy 1741 SAPPGDPKPEYDLELITSCSSNVSVAHADSGKRVYVLTTRDPTTPLARAAWETASHTPVN 1800  
Db 1741 SAPPGDPKPEYDLELITSCSSNVSVAHADSGKRVYVLTTRDPTTPLARAAWETASHTPVN 1800  
Qy 1801 SWLGNIIWYATLWARMLIMTHFFSILLAOBLEKALDCQIYGACYSIEPDLPOIIOQL 1860  
Db 1801 SWLGNIIWYATLWARMLIMTHFFSILLAOBLEKALDCQIYGACYSIEPDLPOIIOQL 1860  
Qy 1861 HGLSAPSLHSYSPEINRVASCLRKLGVPPLRVVHRHARSVRARLLSOGRAATCGKYLIF 1920  
Db 1861 HGLSAPSLHSYSPEINRVASCLRKLGVPPLRVVHRHARSVRARLLSOGRAATCGKYLIF 1920  
Qy 1921 NNAVRTKLKLTPIPAASQDLSSMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIY 1980  
Db 1921 NNAVRTKLKLTPIPAASQDLSSMFVAGYSGGDIYHLSLRARPRFWMCLLLLSVGVGIY 1980  
Qy 1981 LLPNR 1985  
Db 1981 LLPNR 1985

## RESULT 2

US-09-539-601-12  
; Sequence 12, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 12  
; LENGTH: 1985  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-12

Query Match 100.0%; Score 10465; DB 4; Length 1985;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1985; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPITAYSQQTRGLLGCIIITSLTGRDRNQVGEVQVWSTATQSFATCVNGVCWTVYHGA 60  
Db 1 MAPITAYSQQTRGLLGCIIITSLTGRDRNQVGEVQVWSTATQSFATCVNGVCWTVYHGA 60  
Qy 61 GSKTLAGPKGPIQWYTNVDQDLVQWQAPPQARSILTPCTCGSSDLYLVTIRHADVIPVRRR 120  
Db 61 GSKTLAGPKGPIQWYTNVDQDLVQWQAPPQARSILTPCTCGSSDLYLVTIRHADVIPVRRR 120  
Qy 121 GDSRGSLLSPRPVSVLYKSGSGGLLCPSGHVGIFPRAAVCTRGVAKAVDFVPEVSEMETTM 180  
Db 121 GDSRGSLLSPRPVSVLYKSGSGGLLCPSGHVGIFPRAAVCTRGVAKAVDFVPEVSEMETTM 180

181 RSPVFTDNSSPPAVPOTFOVAHLHAPITGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 240  
 181 RSPVFTDNSSPPAVPOTFOVAHLHAPITGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 240  
 241 AYMSKAHGDIPNIRGTGRTITGAPITSTYGVKFLADGGCGGAYDIIICDECHSTSTT 300  
 241 AYMSKAHGDIPNIRGTGRTITGAPITSTYGVKFLADGGCGGAYDIIICDECHSTSTT 300  
 301 ILGIGTVDQAEATAGARLVLATATPGSVTPHPNIEEVALSTGRIPIFYKGAIPETI 360  
 301 ILGIGTVDQAEATAGARLVLATATPGSVTPHPNIEEVALSTGRIPIFYKGAIPETI 360  
 361 KGGHLLIFCHSKKCCDELAALSGLGNNAVAYRGLDVSVIPTSGLDVIWATDALMTGFT 420  
 361 KGGHLLIFCHSKKCCDELAALSGLGNNAVAYRGLDVSVIPTSGLDVIWATDALMTGFT 420  
 421 GDFSVIDCNCVTQTVDPSLDPTFTIETTPQDAYRSORRGRTGGRMGIIYRFVTPG 480  
 421 GDFSVIDCNCVTQTVDPSLDPTFTIETTPQDAYRSORRGRTGGRMGIIYRFVTPG 480  
 481 ERPSGMPDSSVLCBXYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWESVFTGL 540  
 481 ERPSGMPDSSVLCBXYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWESVFTGL 540  
 541 THIDAHFLSQTQAGDNPPYLVAQVATVCARAQAPPPSWDQWKKLIRLKTPLHGTPL 600  
 541 THIDAHFLSQTQAGDNPPYLVAQVATVCARAQAPPPSWDQWKKLIRLKTPLHGTPL 600  
 601 YRLGAVQNEVTTTHITKYIMACHSADLEVTSTWLVGGVLAALAAAYCLTGSVVIYGR 660  
 601 YRLGAVQNEVTTTHITKYIMACHSADLEVTSTWLVGGVLAALAAAYCLTGSVVIYGR 660  
 661 IILSKKPAIIPDREVLVREFDEMBCASHLPYIEQGMOLAEQFKOAKGLQATKQAEA 720  
 661 IILSKKPAIIPDREVLVREFDEMBCASHLPYIEQGMOLAEQFKOAKGLQATKQAEA 720  
 721 AAPVVESKWRITLFAFWAKHMMNFIISGQYLAGSLTLPGNPAIASLMAFTASITSLTQH 780  
 721 AAPVVESKWRITLFAFWAKHMMNFIISGQYLAGSLTLPGNPAIASLMAFTASITSLTQH 780  
 781 TLLFNILGWSAAQALAPSASAFVAGIAGAAVSGIGLKVLDILAGYGAGVAGALVA 840  
 781 TLLFNILGWSAAQALAPSASAFVAGIAGAAVSGIGLKVLDILAGYGAGVAGALVA 840  
 841 FKVMGEMPESTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGECAVQWNNLLAFASR 900  
 841 FKVMGEMPESTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGECAVQWNNLLAFASR 900  
 901 GNHVSPTHVVPESDAAARVTQILSLTITQLLKLHQMINEDCSTPCSGSWLRDVMWDMIC 960  
 901 GNHVSPTHVVPESDAAARVTQILSLTITQLLKLHQMINEDCSTPCSGSWLRDVMWDMIC 960  
 961 TVLTDFKTLWQSKLPLRLCPVPFPFCORGYKGVWRGDMOTTCPCGAOITGHVKNXGMR 1020  
 961 TVLTDFKTLWQSKLPLRLCPVPFPFCORGYKGVWRGDMOTTCPCGAOITGHVKNXGMR 1020  
 1021 IVGPRTCSTNWTGTPINAYTTGCTPPSPAPNVSRALMEVAAEYVEVTRVGDHFVVTGM 1080  
 1021 IVGPRTCSTNWTGTPINAYTTGCTPPSPAPNVSRALMEVAAEYVEVTRVGDHFVVTGM 1080  
 1081 TTDNVKPCQVPAPEFTEVDGVRLLHRYAPACKPLLRBEVTFVLGNQVLVGSQLPCEPE 1140  
 1081 TTDNVKPCQVPAPEFTEVDGVRLLHRYAPACKPLLRBEVTFVLGNQVLVGSQLPCEPE 1140  
 1141 PDVAVLTSMLTDPESHITAEAKRLARGSPPSLASSASQLSAPSKATCTTHDSDPAD 1200  
 1141 PDVAVLTSMLTDPESHITAEAKRLARGSPPSLASSASQLSAPSKATCTTHDSDPAD 1200  
 1201 LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSVAEILRRSRKFPFRAM 1260  
 1201 LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSVAEILRRSRKFPFRAM 1260

1261 PIWADYNPPLLESKMDPDYVPPVHGCPLPPAKAPPIPPPRKRTVWLSSEVSALA 1320  
 1261 PIWADYNPPLLESKMDPDYVPPVHGCPLPPAKAPPIPPPRKRTVWLSSEVSALA 1320  
 1321 ELATKTFGSSSESAVDSGTATASPDQPSDDGAGSDVESYSSMPPLGEPEGDPDLSDGSW 1380  
 1321 ELATKTFGSSSESAVDSGTATASPDQPSDDGAGSDVESYSSMPPLGEPEGDPDLSDGSW 1380  
 1381 STVSEASESDVCCMSYTWGALITPCAAEETKLPINALNSLLRHNLVYATTSKAS 1440  
 1381 STVSEASESDVCCMSYTWGALITPCAAEETKLPINALNSLLRHNLVYATTSKAS 1440  
 1441 LRQKVTDFRLQVLDHRYDVLKEMKASTVAKLLSVEEACKLTTPHSARSFGYGAK 1500  
 1441 LRQKVTDFRLQVLDHRYDVLKEMKASTVAKLLSVEEACKLTTPHSARSFGYGAK 1500  
 1501 DVRNLSSKANVHRSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLLVFPDLG 1560  
 1501 DVRNLSSKANVHRSVWKDLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLLVFPDLG 1560  
 1561 VRVCEKMALVDVSTLPAQVNGSSYGFQYSPGQVFEFLVNAWKAKKCPMGFPAYDTRCFDS 1620  
 1561 VRVCEKMALVDVSTLPAQVNGSSYGFQYSPGQVFEFLVNAWKAKKCPMGFPAYDTRCFDS 1620  
 1621 TVTENDIRVBESIYQCCDLAPEARQALRSALTERLYIGGPLTNSKGQNGCYRRCRASGVLT 1680  
 1621 TVTENDIRVBESIYQCCDLAPEARQALRSALTERLYIGGPLTNSKGQNGCYRRCRASGVLT 1680  
 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY 1740  
 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY 1740  
 1741 SAPGDPKPKBYDELITSCSSNVSAHADSGKAVYLTROPTTPLARAAMETARHTFPV 1800  
 1741 SAPGDPKPKBYDELITSCSSNVSAHADSGKAVYLTROPTTPLARAAMETARHTFPV 1800  
 1801 SWLGNITWYAPTLWARMILMTHFTSILLAOLEKALDCQIYGACYSIEPLDLP01I0RL 1860  
 1801 SWLGNITWYAPTLWARMILMTHFTSILLAOLEKALDCQIYGACYSIEPLDLP01I0RL 1860  
 1861 HGLSAFSLHSPGEBINRVASCLRLKGVPLRVWHRARSVRARLLSOGGRAATCGKYLIF 1920  
 1861 HGLSAFSLHSPGEBINRVASCLRLKGVPLRVWHRARSVRARLLSOGGRAATCGKYLIF 1920  
 1921 NWAVTKLKLTPIPAASQDLSSWFAVSGGDIYHSLSRARPRWFMCLLLLSVGVGLY 1980  
 1921 NWAVTKLKLTPIPAASQDLSSWFAVSGGDIYHSLSRARPRWFMCLLLLSVGVGLY 1980  
 1981 LLENR 1985  
 1981 LLENR 1985

RESULT 3  
 US-09-539-601-6  
 ; Sequence 6, Application US/09539601C  
 ; Patent No. 6630343  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bartenschlager, Ralf FW  
 ; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
 ; FILE REFERENCE: all sequences  
 ; CURRENT APPLICATION NUMBER: US/09/539,601C  
 ; CURRENT FILING DATE: 2001-08-30  
 ; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
 ; EARLIER FILING DATE: 1999-04-03  
 ; NUMBER OF SEQ ID NOS: 51  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 6  
 ; LENGTH: 2201  
 ; TYPE: PRS  
 ; ORGANISM: Hepatitis C virus  
 US-09-539-601-6

Query Match 100.0%; Score 10462; DB 4; Length 2201;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPTAYSOQTRGLLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCMVYHGA 60  
DB 217 LAPITAYSOQTRGLLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCMVYHGA 276

QY 61 GSKTLAGPKGPIOTMYTNVDDQVLGWAQPPGARSITPCTCGSSDLVLYTRHADVIVRRR 120  
DB 277 GSKTLAGPKGPIOTMYTNVDDQVLGWAQPPGARSITPCTCGSSDLVLYTRHADVIVRRR 336

QY 121 GDSRGSLLSPRPVSYLKGSSGCLLPCSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM 180  
DB 337 GDSRGSLLSPRPVSYLKGSSGCLLPCSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM 396

QY 181 RSPVFTONSPPAVPOTFOVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 240  
DB 397 RSPVFTONSPPAVPOTFOVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 456

QY 241 AYMSKAHGIDPNIRGTITGAPITVSTYKFLADGGCGGAYDIIICDECHSTDTST 300  
DB 457 AYMSKAHGIDPNIRGTITGAPITVSTYKFLADGGCGGAYDIIICDECHSTDTST 516

QY 301 ILGIGTVLDOAETAGARLVLATATPPGVSVPVPHNIEEVALSSTGEBIPFYKAIPIETI 360  
DB 517 ILGIGTVLDOAETAGARLVLATATPPGVSVPVPHNIEEVALSSTGEBIPFYKAIPIETI 576

QY 361 KGGHLLFCCHSKKCKDELAKLGLGNAYAYRGLDVSIPTSGDVIVVATDALMTGFT 420  
DB 577 KGGHLLFCCHSKKCKDELAKLGLGNAYAYRGLDVSIPTSGDVIVVATDALMTGFT 636

QY 421 GDFSVIDCNTCVTQTQVDFSLDPTFTTETTPQDAVRSORRGRTGRGMGIYRFVTPG 480  
DB 637 GDFSVIDCNTCVTQTQVDFSLDPTFTTETTPQDAVRSORRGRTGRGMGIYRFVTPG 696

QY 481 ERPSGMFDSVLCEDYDAGCAWYELTPAETSRLRAYLNTPLPVCQDHLFEWESVPTGL 540  
DB 697 ERPSGMFDSVLCEDYDAGCAWYELTPAETSRLRAYLNTPLPVCQDHLFEWESVPTGL 756

QY 541 THIDAHFLSKQAGNFPYLVAYQATVCAQAAPPSPDQMKCLIRLKTPLHGTPTLL 600  
DB 757 THIDAHFLSKQAGNFPYLVAYQATVCAQAAPPSPDQMKCLIRLKTPLHGTPTLL 816

QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVVSTWLVGVLAAALAYCLTTGSSVIVGR 660  
DB 817 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVSTWLVGVLAAALAYCLTTGSSVIVGR 876

QY 661 IILSGKPAIIPDREVLVYRFBDEMECASHLPYIEQGMQLAEQPKQKAIQLQATKQAEA 720  
DB 877 IILSGKPAIIPDREVLVYRFBDEMECASHLPYIEQGMQLAEQPKQKAIQLQATKQAEA 936

QY 721 AAPVVEKWTLEAFWAKHWNPIFSGIYVLAGLSTLPGNPATIASLMAFTASITSPITTOH 780  
DB 937 AAPVVEKWTLEAFWAKHWNPIFSGIYVLAGLSTLPGNPATIASLMAFTASITSPITTOH 996

QY 781 TLLFNILGCVAAQLAPPSSAFAVGAGIAGAVSGIGLKVLDVILAGYGVAGALVA 840  
DB 997 TLLFNILGCVAAQLAPPSSAFAVGAGIAGAVSGIGLKVLDVILAGYGVAGALVA 1056

QY 841 FKVMSEMPSTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGEVQVMMNRLTAFASR 900  
DB 1057 FKVMSEMPSTEDLVNLLPAIILSPGALVGVVCAAILRRHVGPGEVQVMMNRLTAFASR 1116

QY 901 GNEVSTHYVPSDAAARVQILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 960  
DB 1117 GNEVSTHYVPSDAAARVQILSSLTITQLKRLHOWINEDCSTPCSGSMLRDVMDWIC 1176

QY 961 TVLTDFKTLQSKLLPRLPGVFPFSCQYKGVWBGDGMQTTCPGQAQITGHVKNXGMR 1020  
DB 1177 TVLTDFKTLQSKLLPRLPGVFPFSCQYKGVWBGDGMQTTCPGQAQITGHVKNXGMR 1236

QY 1021 IVGPRCTSNWHTGCTPFINAYTTGCTPSPAPNYSRALMRVAAEEYVEVTRVGDHFVVTGM 1080

RESULT 4

US-09-539-601-15  
; Sequence 15, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:

DB 1237 IVGPRCTSNWHTGCTPFINAYTTGCTPSPAPNYSRALMRVAAEEYVEVTRVGDHFVVTGM 1296

QY 1081 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPILLREVTFLVGLNOYLVSQJPCBPE 1140

DB 1297 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPILLREVTFLVGLNOYLVSQJPCBPE 1356

QY 1141 PDVAVLTSMLTDPDPSHIITAEATKRLARGSPPSLASSASQSLKATCTTRHDSPAD 1200

DB 1357 PDVAVLTSMLTDPDPSHIITAEATKRLARGSPPSLASSASQSLKATCTTRHDSPAD 1416

QY 1201 LIEANLLWRQSMGNGNITRVESENKVILDSPEPLQAEDEDEVEVPAEILRRSRKFPFRAM 1260

DB 1417 LIEANLLWRQSMGNGNITRVESENKVILDSPEPLQAEDEDEVEVPAEILRRSRKFPFRAM 1476

QY 1261 PIWARPDPNPLLESWKDPDYPVPPVHGCPLPPAKAPPIPPRRKRTVILSESVSSALA 1320

DB 1477 PIWARPDPNPLLESWKDPDYPVPPVHGCPLPPAKAPPIPPRRKRTVILSESVSSALA 1536

QY 1321 ELATKTFGSSSSAVDSGTATASPDQSDDGADSDVESYSMPPLSEGPDPDLSGWS 1380

DB 1537 ELATKTFGSSSSAVDSGTATASPDQSDDGADSDVESYSMPPLSEGPDPDLSGWS 1596

QY 1381 STVSEEAESDVCCMSYTWGALITPCAEEETKLPIINALSNSLLRHNLVYATTSRSAS 1440

DB 1597 STVSEEAESDVCCMSYTWGALITPCAEEETKLPIINALSNSLLRHNLVYATTSRSAS 1656

QY 1441 LRQKVTDFRLQVLDHHRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFGYCAK 1500

DB 1657 LRQKVTDFRLQVLDHHRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFGYCAK 1716

QY 1501 DVNRLSKAVNHRSVWKKLLEDETEPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560

DB 1717 DVNRLSKAVNHRSVWKKLLEDETEPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1776

QY 1561 VRVCEKMALVDVYSTLPOAVMGSSYFQYSPGQVFLVNAWKAKCPMGFAYDTRCFDS 1620

DB 1777 VRVCEKMALVDVYSTLPOAVMGSSYFQYSPGQVFLVNAWKAKCPMGFAYDTRCFDS 1836

QY 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQVYRRCRASGVL 1680

DB 1837 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQVYRRCRASGVL 1896

QY 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICBSAGTQDEASLRAFTEAMTRY 1740

DB 1897 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICBSAGTQDEASLRAFTEAMTRY 1956

QY 1741 SAPPDGPPEVDELELITSCSSNVSAHDASGKRVYLTREDPTPLARAAMETARHTPVN 1800

DB 1957 SAPPDGPPEVDELELITSCSSNVSAHDASGKRVYLTREDPTPLARAAMETARHTPVN 2016

QY 1801 SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIQRL 1860

DB 2017 SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIQRL 2076

QY 1861 HGLSAPLSHSYSGPEINRVASCLRLGVPPLRVWRHARSVRARLLSQSGRAATCKYLF 1920

DB 2077 HGLSAPLSHSYSGPEINRVASCLRLGVPPLRVWRHARSVRARLLSQSGRAATCKYLF 2136

QY 1921 NNAVRTKLLTPIPAASQDLSSWFVAGYSGGDIYHLSLRARPRFWMWCLLLLSVGVGIY 1980

DB 2137 NNAVRTKLLTPIPAASQDLSSWFVAGYSGGDIYHLSLRARPRFWMWCLLLLSVGVGIY 2196

QY 1981 LLPNR 1985

DB 2197 LLPNR 2201

```
; APPLICANT: Bartschlag, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; EARLIER FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 15
; LENGTH: 2201
; TYPE: PR
; ORGANISM: Hepatitis C virus
; US-09-539-601-15

Query Match 100.0%; Score 10462; DB 4; Length 2201;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MAPITAYSOQTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCNGVCMVTVVHGA 60
: |||||
Db 217 LAPITAYSOQTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCNGVCMVTVVHGA 276

Qy 61 GSKTLAGEKGPITOMYTNVDODLVGQWQAPPGARSLTPTCTCGSSDLVLTTRHADVIPVRR 120
: |||||
Db 277 GSKTLAGEKGPITOMYTNVDODLVGQWQAPPGARSLTPTCTCGSSDLVLTTRHADVIPVRR 336

Qy 121 GDSRGSLLSPRPVSYLKGSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVESMETTM 180
: |||||
Db 337 GDSRGSLLSPRPVSYLKGSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVESMETTM 396

Qy 181 RSPVFTDNSSPPAVPQTFQVAHLHAPGSGKSTKVPAAAYAGQKVLVLPNSVAATLGF 240
: |||||
Db 397 RSPVFTDNSSPPAVPQTFQVAHLHAPGSGKSTKVPAAAYAGQKVLVLPNSVAATLGF 456

Qy 241 AYMSKAGHIDNIRGTITGAPITVSTYTKFLADGCGSGGAYDIIICDECHSTSTT 300
: |||||
Db 457 AYMSKAGHIDNIRGTITGAPITVSTYTKFLADGCGSGGAYDIIICDECHSTSTT 516

Qy 301 ILGTGTVLDAQETAGARLVLATATPPGSVTVPHNIEVALSSTGEIPFYKAIPIETI 360
: |||||
Db 517 ILGTGTVLDAQETAGARLVLATATPPGSVTVPHNIEVALSSTGEIPFYKAIPIETI 576

Qy 361 KGGRHILFCHSKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 420
: |||||
Db 577 KGGRHILFCHSKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 636

Qy 421 GDFDSVIDCNTVCTQTVDFSLDPTFTTETTVPODAVSRSORRGRTGGRMGIIYRFVTPG 480
: |||||
Db 637 GDFDSVIDCNTVCTQTVDFSLDPTFTTETTVPODAVSRSORRGRTGGRMGIIYRFVTPG 696

Qy 481 ERPSGMEFDSVLCBICYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWESVFTGL 540
: |||||
Db 697 ERPSGMEFDSVLCBICYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWESVFTGL 756

Qy 541 THIDAHFLSQTKQAGDNPYLVAVOATVCARAQAPPPSDQMWKCLIRLKPITLHGPTPL 600
: |||||
Db 757 THIDAHFLSQTKQAGDNPYLVAVOATVCARAQAPPPSDQMWKCLIRLKPITLHGPTPL 816

Qy 601 YRLGAVQNEVTTTHPIITKYINWACMSADLEVTVSTWLVGGVLAALAAVCLTTGTVVIVGR 660
: |||||
Db 817 YRLGAVQNEVTTTHPIITKYINWACMSADLEVTVSTWLVGGVLAALAAVCLTTGTVVIVGR 876

Qy 661 IILSGKPAIIPDREVLVYREFDEMEECASHLPYIEQGMQLAEQFKQKAIIGLLQTATKQAEA 720
: |||||
Db 877 IILSGKPAIIPDREVLVYREFDEMEECASHLPYIEQGMQLAEQFKQKAIIGLLQTATKQAEA 936

Qy 721 AAPVVESKRWLEAPWAKHMFNTSGIYLAGLSTLPGNPAIASMAFTASITSPITQ 780
: |||||
Db 937 AAPVVESKRWLEAPWAKHMFNTSGIYLAGLSTLPGNPAIASMAFTASITSPITQ 996

Qy 781 TLLFNILGGWAAQLAPPSAASAFVAGIAGAAVGSIGLGVKLVLDILAGYGAGVAGALVA 840
: |||||
Db 2077 HGLSAFSLHSYSGPEINRVASCLKLGVPPLRVWRHARSVRARLLSQGGRAATCGKXLF 2136
```

QY 1921 NWAURTKLKTPIPAASQLDLSWFEVAGYSGGDIYHLSRARPWFWMKCLLLSVGVGIY 1380  
DB 2137 NWAURTKLKTPIPAASQLDLSWFEVAGYSGGDIYHLSRARPWFWMKCLLLSVGVGIY 2196

QY 1981 LLENR 1985  
DB 2197 LLENR 2201

RESULT 5  
US-09-539-601-3  
; Sequence 3, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartschlagel, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 3  
; LENGTH: 3010  
; TYPE: PRT  
; ORGANISM: Hepatitis C virus  
US-09-539-601-3

Query Match 100.0%; Score 10462; DB 4; Length 3010;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLCIIITSLTRDRNQVEGEVQVVSSTATQSFATCVNGVCWTVYHGA 60  
DB 1026 LAPITAYSQOTRGLLCIIITSLTRDRNQVEGEVQVVSSTATQSFATCVNGVCWTVYHGA 1085  
QY 61 GSKTLAGPKGPIQMTYNDVDLVGWQAPPGASRLTPCTGSSDLVYLRHADVIPVRRR 120  
DB 1086 GSKTLAGPKGPIQMTYNDVDLVGWQAPPGASRLTPCTGSSDLVYLRHADVIPVRRR 1145  
QY 121 GDSRGLLPSPVSYLKGSSGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM 180  
DB 1146 GDSRGLLPSPVSYLKGSSGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM 1205  
QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 240  
DB 1206 RSPVFTDNSPPAVPQTFQVAHLHAPTGSCKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 1265  
QY 241 AYMSKAHGIDPNIRTVRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 300  
DB 1266 AYMSKAHGIDPNIRTVRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 1325  
QY 301 ILGIGTVLQDAETAGARLVVLAATATPGSVTVPHPNIEEVALSSTGEIPIYKAIPIETI 360  
DB 1326 ILGIGTVLQDAETAGARLVVLAATATPGSVTVPHPNIEEVALSSTGEIPIYKAIPIETI 1385  
QY 361 KGRGHLIFCHSKKKCKDELAAGLGLGNVAYYRGLDVSVIPTSGDVIIVATDALTMTGFT 420  
DB 1386 KGRGHLIFCHSKKKCKDELAAGLGLGNVAYYRGLDVSVIPTSGDVIIVATDALTMTGFT 1445  
QY 421 GDFSDVIDCNTQVQVDFSLDPTFTIETTTPQDAVSRSORRGTRGRMGRIYRFTPG 480  
DB 1446 GDFSDVIDCNTQVQVDFSLDPTFTIETTTPQDAVSRSORRGTRGRMGRIYRFTPG 1505  
QY 481 ERPSGMPDSSVLCCEYDAGCAWYELTPAETSRLRAYLNTPLPVCODHLEFWESVFTGL 540  
DB 1506 ERPSGMPDSSVLCCEYDAGCAWYELTPAETSRLRAYLNTPLPVCODHLEFWESVFTGL 1565  
QY 541 THIDAFLSQTQAGDNEPVLVAYQATVCARAQAPPPSDMKCLIRLKPTLHGPTPL 600  
DB 1566 THIDAFLSQTQAGDNEPVLVAYQATVCARAQAPPPSDMKCLIRLKPTLHGPTPL 1625

QY 601 YRLGAVQNEVTTHTPIITKYIMACMSADLEVVVTWLVGGVLAALAAAYCLTTSVWIVGR 660  
DB 1626 YRLGAVQNEVTTHTPIITKYIMACMSADLEVVVTWLVGGVLAALAAAYCLTTSVWIVGR 1685  
QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLAEQFKQKAIQLQATKQAEA 720  
DB 1686 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLAEQFKQKAIQLQATKQAEA 1745  
QY 721 AAPVSESKWRTLEAFWAKHWNFIISGIYIAGLSTLPGNPAIASLMAFTASITSLTQH 780  
DB 1746 AAPVSESKWRTLEAFWAKHWNFIISGIYIAGLSTLPGNPAIASLMAFTASITSLTQH 1805  
QY 781 TLLFNILGGWVAQAAPPAAAFVAGAGIAGVSGIIGLKVLDIILAGVAGVAGALVA 840  
DB 1806 TLLFNILGGWVAQAAPPAAAFVAGAGIAGVSGIIGLKVLDIILAGVAGVAGALVA 1865  
QY 841 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQMMNLLIAPASR 900  
DB 1866 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQMMNLLIAPASR 1925  
QY 901 GNVSPHVVPSDAAARVTOILLSSLTITQLLKRHOWINEDCSTPCSGSWLRDWDWIC 960  
DB 1926 GNVSPHVVPSDAAARVTOILLSSLTITQLLKRHOWINEDCSTPCSGSWLRDWDWIC 1985  
QY 961 TVLTDFKTWLOSLLPRLPGVPPFCQGVYKVMRGDGMOTTCPCGAQITGHVKGSMR 1020  
DB 1986 TVLTDFKTWLOSLLPRLPGVPPFCQGVYKVMRGDGMOTTCPCGAQITGHVKGSMR 2045  
QY 1021 IVPRTCSNTWHTGTFPINAITYTGPCTSPAPNYSRALMRVAABEYVEVTRVGFHYVTGM 1080  
DB 2046 IVPRTCSNTWHTGTFPINAITYTGPCTSPAPNYSRALMRVAABEYVEVTRVGFHYVTGM 2105  
QY 1081 TTDNVKCPQVPAPEPTEVDGVLHRYAPACKPILREEVTFVLGNOYLVGSQLPCEPE 1140  
DB 2106 TTDNVKCPQVPAPEPTEVDGVLHRYAPACKPILREEVTFVLGNOYLVGSQLPCEPE 2165  
QY 1141 PDVAVLTSMLTDSHITAEAKRLARGPPSPASSASOLSAPSLKATCTTRHDSPPAD 1200  
DB 2166 PDVAVLTSMLTDSHITAEAKRLARGPPSPASSASOLSAPSLKATCTTRHDSPPAD 2225  
QY 1201 LIEANLLMRQEMGNITRVESENKVLDSFEPLQAEDEEREVSFPAELRRSRKPPRAM 1260  
DB 2226 LIEANLLMRQEMGNITRVESENKVLDSFEPLQAEDEEREVSFPAELRRSRKPPRAM 2285  
QY 1261 PIWARPDPNPPLESKWDYDYPVVGCPPLPAKAPPPIPPRRKKTIVLSESTVSALA 1320  
DB 2286 PIWARPDPNPPLESKWDYDYPVVGCPPLPAKAPPPIPPRRKKTIVLSESTVSALA 2345  
QY 1321 ELATKTFGSESSAVDSGTATASPDQSDDDAGSDVESYSSMPPLEGEPEGDPDLSGWS 1380  
DB 2346 ELATKTFGSESSAVDSGTATASPDQSDDDAGSDVESYSSMPPLEGEPEGDPDLSGWS 2405  
QY 1381 STVSEASBDDVCCMSYTWGTALITPCAABETKLPINALNSLLRHNLLVYATTSRAS 1440  
DB 2406 STVSEASBDDVCCMSYTWGTALITPCAABETKLPINALNSLLRHNLLVYATTSRAS 2465  
QY 1441 LRQKKVTFDLQVLDHVDVLEKMAKASTVYKALLSVEEACKLTPPHSARSKFYGA 1500  
DB 2466 LRQKKVTFDLQVLDHVDVLEKMAKASTVYKALLSVEEACKLTPPHSARSKFYGA 2525  
QY 1501 DVNLSKKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
DB 2526 DVNLSKKAVNHRSVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 2585  
QY 1561 VRVCEKMAIYDVVSTLPOAVMGSSYGFQYSPGQVREFLVNMAKAKCPMGFAVDTKCFDS 1620  
DB 2586 VRVCEKMAIYDVVSTLPOAVMGSSYGFQYSPGQVREFLVNMAKAKCPMGFAVDTKCFDS 2645  
QY 1621 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGLPTNSKQNGCVRRCRASGLT 1680  
DB 2646 TVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGLPTNSKQNGCVRRCRASGLT 2705

QY 1681 TSCGNLTLCYLKAAACAAKLDQCTMLVCGDDLVVLCESAGTOEDASLRAFTAMTRY 1740  
Db 2706 TSCGNLTLCYLKAAACAAKLDQCTMLVCGDDLVVLCESAGTOEDASLRAFTAMTRY 2765  
QY 1741 SAPGDDPKPYDLELITSCSSNVSAHADSGKRVYVLTTRDPTTFLARAAWETARHTPVN 1800  
Db 2766 SAPGDDPKPYDLELITSCSSNVSAHADSGKRVYVLTTRDPTTFLARAAWETARHTPVN 2825  
QY 1801 SWLGNIIWYAPTILWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLQIIQRL 1860  
Db 2826 SWLGNIIWYAPTILWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLQIIQRL 2885  
QY 1861 HGLSAFSLHSYSPGEINRVASCLKGLVPPRLVRVHRARSVRARLLSOGGAAATCGKYL 1920  
Db 2886 HGLSAFSLHSYSPGEINRVASCLKGLVPPRLVRVHRARSVRARLLSOGGAAATCGKYL 2945  
QY 1921 NWAVRTKLKLTPIPAASOLDLSSWFFVAGYSGGDIYHLSLRARPRFWFMCCLLLSVGVGIY 1980  
Db 2946 NWAVRTKLKLTPIPAASOLDLSSWFFVAGYSGGDIYHLSLRARPRFWFMCCLLLSVGVGIY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

## RESULT 6

US-10-029-907-3  
; Sequence 3, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2201  
; TYPE: PRT  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 882  
; OTHER INFORMATION: Xaa is Lys or Arg  
; NAME/KEY: VARIANT  
; LOCATION: 1489  
; OTHER INFORMATION: Xaa is Leu  
US-10-029-907-3

Query Match 99.8%; Score 10442; DB 4; Length 2201;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1981; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MAPITAYSQQRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVCVGVCMVYHGA 60  
Db 217 LAPITAYSQQRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVCVGVCMVYHGA 276  
QY 61 GSKTLAGKGPITQMYTNVDQDLVGWQAPPGARSITPCTCGSSDLVYTRHADVIPVRR 120  
Db 277 GSKTLAGKGPITQMYTNVDQDLVGWQAPPGARSITPCTCGSSDLVYTRHADVIPVRR 336  
QY 121 GDSRGSLLSPRPVSYLKGSSGGLPCPSCHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 337 GDSRGSLLSPRPVSYLKGSSGGLPCPSCHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 396  
QY 181 RSPVFTDNSSPPAVPQTQFVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 240  
Db 397 RSPVFTDNSSPPAVPQTQFVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLGF 456

QY 241 AYMSKAHGIIDPNIITGVRTITTTGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 300  
Db 457 AYMSKAHGIIDPNIITGVRTITTTGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 516  
QY 301 ILGIGTVLDOAETAGARLVLATATPGSVTVPHNPTEEVALSSTGEIPFFYKGAIPETI 360  
Db 517 ILGIGTVLDOAETAGARLVLATATPGSVTVPHNPTEEVALSSTGEIPFFYKGAIPETI 576  
QY 361 KGRHLLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVTSGDVIIVATDALMTGFT 420  
Db 577 KGRHLLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVTSGDVIIVATDALMTGFT 636  
QY 421 GDFSVIDCNTCTVTQTVDFSLDPTFTTITTVPODAVRSORRGRTGRGMGIIRFVTPG 480  
Db 637 GDFSVIDCNTCTVTQTVDFSLDPTFTTITTVPODAVRSORRGRTGRGMGIIRFVTPG 696  
QY 481 ERPSGMFDSVCLCYDAGCAWYELTPAETSRLRAYLNTPLGPVPCQDHLFEWESVFTGL 540  
Db 697 ERPSGMFDSVCLCYDAGCAWYELTPAETSRLRAYLNTPLGPVPCQDHLFEWESVFTGL 756  
QY 541 THIDAHFLSQTQAGDNFPYLVAYOATVCARQAAPPSPSWDOMKCLIRLKTPLHGPTLL 600  
Db 757 THIDAHFLSQTQAGDNFPYLVAYOATVCARQAAPPSPSWDOMKCLIRLKTPLHGPTLL 816  
QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTSGVVI 660  
Db 817 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTGTSGVVI 876  
QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLABQFKOKAIGLLQATKQAEA 720  
Db 877 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLABQFKOKAIGLLQATKQAEA 936  
QY 721 AAPVYESKWRTELEAFWAKHMWNFISGIOYLAGSLTLPENPAIASIMAFASITSPLTQ 780  
Db 937 AAPVYESKWRTELEAFWAKHMWNFISGIOYLAGSLTLPENPAIASIMAFASITSPLTQ 996  
QY 781 TLLFNILGGWVAQALAPPSAASAEVAGIAGAAVSGTIGLKVIVDIIAGYAGAGALVA 840  
Db 997 TLLFNILGGWVAQALAPPSAASAEVAGIAGAAVSGTIGLKVIVDIIAGYAGAGALVA 1056  
QY 841 FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLAFASR 900  
Db 1057 FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLAFASR 1116  
QY 901 GNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHQMINEDEDCSTPCSGSLRDVDMIC 960  
Db 1117 GNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHQMINEDEDCSTPCSGSLRDVDMIC 1176  
QY 961 TVLTDKFTWLSKLLPRLPGVPPFSCORGKYGVRGDMQTTCCGNOITGHVKNCSMR 1020  
Db 1177 TVLTDKFTWLSKLLPRLPGVPPFSCORGKYGVRGDMQTTCCGNOITGHVKNCSMR 1236  
QY 1021 IVGPRTCSNTWHGTFPINAYTTGCTSPAPNYSRALWRVAEEVEVTRVGDHYVTGM 1080  
Db 1237 IVGPRTCSNTWHGTFPINAYTTGCTSPAPNYSRALWRVAEEVEVTRVGDHYVTGM 1296  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVRLHRYAPACKPLLEEVTVLVLNQYLVSQLPCEPE 1140  
Db 1297 TTDNVKCPQVPAPEFFTEVDGVRLHRYAPACKPLLEEVTVLVLNQYLVSQLPCEPE 1356  
QY 1141 PDVAVLTSMLTDPSSHITAETAKRRLARGSPPSLASSASQLSAPSLKATCTTTRHSDPAD 1200  
Db 1357 PDVAVLTSMLTDPSSHITAETAKRRLARGSPPSLASSASQLSAPSLKATCTTTRHSDPAD 1416  
QY 1201 LIEANLLWRQEMGNITRVESENKVILDSFEPLQAEDEREVSVPAEILRRSRKFFRAM 1260  
Db 1417 LIEANLLWRQEMGNITRVESENKVILDSFEPLQAEDEREVSVPAEILRRSRKFFRAM 1476  
QY 1261 PIWARPYNPLLESWKDPDYPVPVHVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 1477 PIWARPYNPLLESWKDPDYPVPVHVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1536  
QY 1321 ELATKTFGSSSESSAVDSGTATASDPQSDDDGSDVESYSSMPPLEGEPCDPLSDGSW 1380



1537 ELATKTFGSSSESAVDSGTATASPDQSDSDGSDVESYSSPPLEGEFGEDLSDGSW 1596  
 1381 STVSEASESDVVCSSMSYTWGALITPCAAABETKLPINALNSLLRHHNLIYYATTSSAS 1440  
 1597 STVSEASESDVVCSSMSYTWGALITPCAAABETKLPINALNSLLRHHNLIYYATTSSAS 1656  
 1441 LROKKTFTDLQVLDHRYDVLKEMKAKASTVAKLLSVEAECKLTPPHSARSKFGYGAK 1500  
 1657 LROKKTFTDLQVLDHRYDVLKEMKAKASTVAKLLSVEAECKLTPPHSARSKFGYGAK 1716  
 1501 DVRLSSKAVNHRSVWKDQLEDETETPIDITIMAKNEVFCVQPEKGGKRPALIVFPDLG 1560  
 1717 DVRLSSKAVNHRSVWKDQLEDETETPIDITIMAKNEVFCVQPEKGGKRPALIVFPDLG 1776  
 1561 VRCEKMALYDVVSTLTPQAVMGSSYGHQSPGQRFVFLVNAWKAKCPMGFAVDTTRCFDS 1620  
 1777 VRCEKMALYDVVSTLTPQAVMGSSYGHQSPGQRFVFLVNAWKAKCPMGFAVDTTRCFDS 1836  
 1621 TVTENDIRVEESYQCCDLAPARQATRSILTERLYTGGPLTNSKGQNGCYRRCRASGVLT 1680  
 1837 TVTENDIRVEESYQCCDLAPARQATRSILTERLYTGGPLTNSKGQNGCYRRCRASGVLT 1896  
 1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLWVICSAGTQDEEASLRAFTEAMTRY 1740  
 1897 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLWVICSAGTQDEEASLRAFTEAMTRY 1956  
 1741 SAPPGPPKPEYDELELITSCSSNVSAHADSGKRVYLLTRDPTPLARAAWETARHTPVN 1800  
 1957 SAPPGPPKPEYDELELITSCSSNVSAHADSGKRVYLLTRDPTPLARAAWETARHTPVN 2016  
 1801 SWLGNIIWAPTILWARMILMTHFFSILLAEQLEKALDCQIYACYSIEPLDLQIILQRL 1860  
 2017 SWLGNIIWAPTILWARMILMTHFFSILLAEQLEKALDCQIYACYSIEPLDLQIILQRL 2076  
 1861 HGLSAPLSHSYSGEINRVASCLRLKGVPLRVHRARSVRARLLSQGGRAATCGKLYF 1920  
 2077 HGLSAPLSHSYSGEINRVASCLRLKGVPLRVHRARSVRARLLSQGGRAATCGKLYF 2136  
 1921 NWAVRTKLTPIPAASQLDLSNFWAGYSGGDIYHLSRPRPFWMCILLISVGVGIY 1980  
 2137 NWAVRTKLTPIPAASQLDLSNFWAGYSGGDIYHLSRPRPFWMCILLISVGVGIY 2196  
 1981 LLENR 1985  
 2197 LLENR 2201

RESULT 7  
 US-09-539-601-24  
 ; Sequence 24, Application US/09539601C  
 ; Patent No. 6630343  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bartschlagel, Ralf FW  
 ; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
 ; FILE REFERENCE: all sequences  
 ; CURRENT APPLICATION NUMBER: US/09/539,601C  
 ; CURRENT FILING DATE: 2001-08-30  
 ; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
 ; EARLIER FILING DATE: 1999-04-03  
 ; NUMBER OF SEQ ID NOS: 51  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 24  
 ; LENGTH: 1985  
 ; TYPE: PRT  
 ; ORGANISM: Hepatitis C virus  
 US-09-539-601-24

Query Match 99.7%; Score 10435; DB 4; Length 1985;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1979; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

1 MAPITAYSQOTRGLGCIITSLTGRDRNVEGEVQVSTATQSFATCVNGVCWTVYHGA 60  
 61 GSKTLGAPKGPITOMTNTVNDQDLVGMQAPPGARSLLTPTCTGSSDLYLVTBHADVIVRRR 120  
 61 GSKTLGAPKGPITOMTNTVNDQDLVGMQAPPGARSLLTPTCTGSSDLYLVTBHADVIVRRR 120  
 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRCVAKAVDFVPVSMETTM 180  
 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRCVAKAVDFVPVSMETTM 180  
 181 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSCKSTKVPAAVAAQGYKVLVLPNSVAATLFG 240  
 181 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSCKSTKVPAAVAAQGYKVLVLPNSVAATLFG 240  
 241 AYNKKAHGDIPNIRITGVRTITTCAPITYSTYGFADGGCGSGGAYDIIICDECHSTDSTT 300  
 241 AYNKKAHGDIPNIRITGVRTITTCAPITYSTYGFADGGCGSGGAYDIIICDECHSTDSTT 300  
 301 ILGIGTVLQDAETAGARLVVLAATATPPGVTVPHPNIEEVALSSTGEIPYGKAIPLETI 360  
 301 ILGIGTVLQDAETAGARLVVLAATATPPGVTVPHPNIEEVALSSTGEIPYGKAIPLETI 360  
 361 KGRHILIFCHSKKKKCDLAALKSLGGLNVAAYYRGLDVSVIPTSGDVIVVATDALMTGFT 420  
 361 KGRHILIFCHSKKKKCDLAALKSLGGLNVAAYYRGLDVSVIPTSGDVIVVATDALMTGFT 420  
 421 GPDFSDVDCNTCVTQVDFSLDPTFTIETTVQDQAVSRQRRGTRGRHMYRFPVTPG 480  
 421 GPDFSDVDCNTCVTQVDFSLDPTFTIETTVQDQAVSRQRRGTRGRHMYRFPVTPG 480  
 481 ERPSGMFDSVLCCEVDAGCAWYELTPAETSVELRAYLNTPLGVPQDHLHFEESVFTGL 540  
 481 ERPSGMFDSVLCCEVDAGCAWYELTPAETSVELRAYLNTPLGVPQDHLHFEESVFTGL 540  
 541 THIDAFLSQTQAGDNFFYLVAQATVCARAQAPPPSWDMQWKCLIRLKPTLHGPPTLL 600  
 541 THIDAFLSQTQAGDNFFYLVAQATVCARAQAPPPSWDMQWKCLIRLKPTLHGPPTLL 600  
 601 YRLGAVQNEVTHPTITTKYIMACMSADLEVVSTWLVGVLAALAAAYCLTTCSSVIVGR 660  
 601 YRLGAVQNEVTHPTITTKYIMACMSADLEVVSTWLVGVLAALAAAYCLTTCSSVIVGR 660  
 661 IILSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEQKQKAIQLQATKQAEA 720  
 661 IILSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEQKQKAIQLQATKQAEA 720  
 721 AAPVVESKWRTEAFWAKHMMNFIISQIYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
 721 AAPVVESKWRTEAFWAKHMMNFIISQIYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
 781 TLLFNILGGWVAQAQLAPPSAASAFVAGTAGAAGVSGILGKVLAVDLIAGVAGVAGALVA 840  
 781 TLLFNILGGWVAQAQLAPPSAASAFVAGTAGAAGVSGILGKVLAVDLIAGVAGVAGALVA 840  
 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQMMNLIASFAS 900  
 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQMMNLIASFAS 900  
 901 GNVHSPTHVPSDAAAARVTQILSSITITOLLKRLHWINEDCSTPCSSGSLRDVWDWIC 960  
 901 GNVHSPTHVPSDAAAARVTQILSSITITOLLKRLHWINEDCSTPCSSGSLRDVWDWIC 960  
 961 TVLTDFKTLQSKLLPRLPGVPFFSCORGKGVWRGDMOTTCPCGAOITGHVKNKSMR 1020  
 961 TVLTDFKTLQSKLLPRLPGVPFFSCORGKGVWRGDMOTTCPCGAOITGHVKNKSMR 1020  
 1021 IGVPRTCSTNTHGTFPINAITYTTCPTSPAPNYSRALWRVAABEYVEVTRGVGFHYVTGM 1080  
 1021 IGVPRTCSTNTHGTFPINAITYTTCPTSPAPNYSRALWRVAABEYVEVTRGVGFHYVTGM 1080  
 1081 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPILREEVTFVLGNQLVGLSOLPCEPE 1140



Db 1081 TTDDVKPCQCPAPEBFTTEVDGVLHRYAPACKPLLEEVTFLVGLNQLVGSQLPCEPE 1140  
QY 1141 PVAVLUTSMLTTPSHITATKRLARGSPPLSSASSASQLSAPSUKATCTTRHDSPPAD 1200  
Db 1141 PVAVLUTSMLTTPSHITATKRLARGSPPLSSASSASQLSAPSUKATCTTRHDSPPAD 1200  
QY 1201 LIEANLLMQEMGNITRYESENKVVILDSFPLQAEEDEREVSVAEILRRSRKPPRAM 1260  
Db 1201 LIEANLLMQEMGNITRYESENKVVILDSFPLQAEEDEREVSVAEILRRSRKPPRAM 1260  
QY 1261 PIWAPDYNNPPLLESWKODPYVPVHGCPLPPAKAPPITPPRRKRTVVLSESTVSSALA 1320  
Db 1261 PIWAPDYNNPPLLESWKODPYVPVHGCPLPPAKAPPITPPRRKRTVVLSESTVSSALA 1320  
QY 1321 ELATITFGSSSESAVDSGTATASPOPSDDGAGSDVESYSMPLEGGPDPDLSDGWS 1380  
Db 1321 ELATITFGSSSESAVDSGTATASPOPSDDGAGSDVESYSMPLEGGPDPDLSDGWS 1380  
QY 1381 STVSSEASEDVVCCSMSYTWTCALITPCAAEETKLPINALSNLLRHHNLVATTSSAS 1440  
Db 1381 STVSSEASEDVVCCSMSYTWTCALITPCAAEETKLPINALSNLLRHHNLVATTSSAS 1440  
QY 1441 LRQKVTDRLOVLDHDDHVDLKENKAKASTVKALLSVEEACKLTPPHSARSKFYGAK 1500  
Db 1441 LRQKVTDRLOVLDHDDHVDLKENKAKASTVKALLSVEEACKLTPPHSARSKFYGAK 1500  
QY 1501 DVRLNSSKAVNHRSVWKDLLEDTEPTDITTIMAKNEVFCVQPEKGRKPARLIYFPDLG 1560  
Db 1501 DVRLNSSKAVNHRSVWKDLLEDTEPTDITTIMAKNEVFCVQPEKGRKPARLIYFPDLG 1560  
QY 1561 VRVCEKMALYDVWSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKCPMGFAYDTRCFDS 1620  
Db 1561 VRVCEKMALYDVWSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKCPMGFAYDTRCFDS 1620  
QY 1621 TVTENDIRVERSIYOCCLDAPPEARQAIRSLTERLYIGGLTNSKQNCGYRRCRASGVLT 1680  
Db 1621 TVTENDIRVERSIYOCCLDAPPEARQAIRSLTERLYIGGLTNSKQNCGYRRCRASGVLT 1680  
QY 1681 TSCGNLTLCYLKAAACRAAKLODCTMLVCGDDLVIICESAGTQDEASLAFTEAMTRY 1740  
Db 1681 TSCGNLTLCYLKAAACRAAKLODCTMLVCGDDLVIICESAGTQDEASLAFTEAMTRY 1740  
QY 1741 SAPPGDPPKPYDLELITSCSSNVSVAHDASGRVYLTTRPTTPPLARAAMETARHTPVN 1800  
Db 1741 SAPPGDPPKPYDLELITSCSSNVSVAHDASGRVYLTTRPTTPPLARAAMETARHTPVN 1800  
QY 1801 SWLGNIIIMYAPTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLQIIQRL 1860  
Db 1801 SWLGNIIIMYAPTLWARMILMTHFFSILLAQBLEKALDCQIYGACYSIEPLDLQIIQRL 1860  
QY 1861 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSOGGAAATCGKYL 1920  
Db 1861 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRVWRHRARSVRARLLSOGGAAATCGKYL 1920  
QY 1921 NWAVRTKLUTPIPAASQDLSSWFVAGYSGGDIYHSLSRARPFWFWMCLLLSVGVGIY 1980  
Db 1921 NWAVRTKLUTPIPAASQDLSSWFVAGYSGGDIYHSLSRARPFWFWMCLLLSVGVGIY 1980  
QY 1981 LLPNR 1985  
Db 1981 LLPNR 1985

RESULT 8  
; Sequence 27, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30

; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 27  
; LENGTH: 3010  
; TYPE: PR1  
; ORGANISM: Hepatitis C virus  
US-09-539-601-27

Query Match 99.7%; Score 10432; DB 4; Length 3010;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 1978; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLGCIIITSITGRDRNOVEGEVQVSTATOSFLATCVNGVCWTYYHGA 60  
Db 1026 LAPITAYSQOTRGLLGCIIITSITGRDRNOVEGEVQVSTATOSFLATCVNGVCWTYYHGA 1085  
QY 61 GSKTLAGPKGPITOMYTNVDQLVGMQAPPAGARSITPCTCGSSDLYLVTRHADVIPVRRR 120  
Db 1086 GSKTLAGPKGPITOMYTNVDQLVGMQAPPAGARSITPCTCGSSDLYLVTRHADVIPVRRR 1145  
QY 121 GDSRGLLSRPVSVYKSSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDFPVESMETTM 180  
Db 1146 GDSRGLLSRPVSVYKSSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDFPVESMETTM 1205  
QY 181 RSPVFTDNSSPPAVPQTEQVAHLHAPTSGSKSTKVPAAAYAAQYKVLNPSVAATLGG 240  
Db 1206 RSPVFTDNSSPPAVPQTEQVAHLHAPTSGSKSTKVPAAAYAAQYKVLNPSVAATLGG 1265  
QY 241 AYMSKAHGIDNIRTVGRTITITGAPITVSTYKFLADGCGSGGAYDIIICDECHSTDSTT 300  
Db 1266 AYMSKAHGIDNIRTVGRTITITGAPITVSTYKFLADGCGSGGAYDIIICDECHSTDSTT 1325  
QY 301 ILGIGTVLDQAEAGARLVLATATPPGSVTVPHNIEEVALSSTGEIPFYKAIPIETI 360  
Db 1326 ILGIGTVLDQAEAGARLVLATATPPGSVTVPHNIEEVALSSTGEIPFYKAIPIETI 1385  
QY 361 KGRHILFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 420  
Db 1386 KGRHILFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 1445  
QY 421 GDFDSVIDCNTCVTQTVDFSLDPTFTTITTTVPQDAVSRSQRRGTRGRMGIYFVTPG 480  
Db 1446 GDFDSVIDCNTCVTQTVDFSLDPTFTTITTTVPQDAVSRSQRRGTRGRMGIYFVTPG 1505  
QY 481 ERPSGMDSSVLCBCYDAGCAWYELTPAETSVRIRAYLNTPLGVPQCQHLBFWESVFTGL 540  
Db 1506 ERPSGMDSSVLCBCYDAGCAWYELTPAETSVRIRAYLNTPLGVPQCQHLBFWESVFTGL 1565  
QY 541 THIDAHFLSQTQKAGDNFPYLVAYQATVCARAQAPPSPSWDMWKCLRLKPTLHGPTPL 600  
Db 1566 THIDAHFLSQTQKAGDNFPYLVAYQATVCARAQAPPSPSWDMWKCLRLKPTLHGPTPL 1625  
QY 601 YRLGAVONEVTTTHPIYKIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSAVIVGR 660  
Db 1626 YRLGAVONEVTTTHPIYKIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSAVIVGR 1685  
QY 661 IILSGKPAIIPDREVLRYEDEMEECASHLPYIEQGMQLAEQFKQKAILLOTATKQAEA 720  
Db 1686 IILSGKPAIIPDREVLRYEDEMEECASHLPYIEQGMQLAEQFKQKAILLOTATKQAEA 1745  
QY 721 AAPVYESKWTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780  
Db 1746 AAPVYESKWTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTQH 1805  
QY 781 TLLFNILGGWVAQAQLAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA 840  
Db 1806 TLLFNILGGWVAQAQLAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA 1865  
QY 841 FKVMSGEMSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQMMNLIIFASR 900  
Db 1866 FKVMSGEMSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQMMNLIIFASR 1925

QY	901	GNVSPFTHVPSDAAARVTQILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC	960
Db	1926	GNHVSPTHVPSDAAARVTQILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC	1985
QY	961	TVLTDFKTLQSKLLRLPGVPFSCORGKGVWVRGDMQTCPCGAQITGHVKNKSMR	1020
Db	1986	TVLTDFKTLQSKLLRLPGVPFSCORGKGVWVRGDMQTCPCGAQITGHVKNKSMR	2045
QY	1021	IVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDVHYVTGM	1080
Db	2046	IVGPRTCSNTHGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGDVHYVTGM	2105
QY	1081	TTDNVCKPCQVPAPEFTEVDGVRHLRYAPACKPLLEBEVTVFLVGLNOYLVSQLPCEPE	1140
Db	2106	TTDNVCKPCQVPAPEFTEVDGVRHLRYAPACKPLLEBEVTVFLVGLNOYLVSQLPCEPE	2165
QY	1141	PDVAVLTSMLTDPDSHTAETAKRLARGSPPLASSASQLSAPSLKATCTTRHDSFPDAD	1200
Db	2166	PDVAVLTSMLTDPDSHTAETAKRLARGSPPLASSASQLSAPSLKATCTTRHDSFPDAD	2225
QY	1201	LIEANLWRQEMGNITRVSEKNKVILDSFEPLQAEEDEREVSVPAAEILRRSRKPPRAM	1260
Db	2226	LIEANLWRQEMGNITRVSEKNKVILDSFEPLQAEEDEREVSVPAAEILRRSRKPPRAM	2285
QY	1261	PIWARDPNPLLESKMDPDYVPPVHGCPLPAKAPPPIPPRRKRTVVLSESTVSALA	1320
Db	2286	PIWARDPNPLLESKMDPDYVPPVHGCPLPAKAPPPIPPRRKRTVVLSESTVSALA	2345
QY	1321	ELATKTFGSSSAVDSGTATASPDQSDGSDGSDVESYSSMPPLGEPEGPDPLSDGSW	1380
Db	2346	ELATKTFGSSSAVDSGTATASPDQSDGSDGSDVESYSSMPPLGEPEGPDPLSDGSW	2405
QY	1381	STVSEEBASEDVCCSMSTYTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSAS	1440
Db	2406	STVSEEBASEDVCCSMSTYTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSRSAS	2465
QY	1441	LQKKVTFDLQVLDDHYRDLVKEMKAKASTVKAKLLSVEEACKLTPPHSARSFEGYAK	1500
Db	2466	LQKKVTFDLQVLDDHYRDLVKEMKAKASTVKAKLLSVEEACKLTPPHSARSFEGYAK	2525
QY	1501	DVRNLSSKANVNHISVWKDLEDETPTDITTIMAKNEVFVCPQPKGGRKPARLLVFPDLG	1560
Db	2526	DVRNLSSKANVNHISVWKDLEDETPTDITTIMAKNEVFVCPQPKGGRKPARLLVFPDLG	2585
QY	1561	VRVCEKVALYDVVSTLPQAVMGSSYFGYSPGQRFVEFLVNAWAKKCPMGPAYDTRCFDS	1620
Db	2586	VRVCEKVALYDVVSTLPQAVMGSSYFGYSPGQRFVEFLVNAWAKKCPMGPAYDTRCFDS	2645
QY	1621	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGPLTNSKGONCGYRRCRASGLT	1680
Db	2646	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGPLTNSKGONCGYRRCRASGLT	2705
QY	1681	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLWVICESAGTOEDASLRAFTTEAMTRY	1740
Db	2706	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLWVICESAGTOEDASLRAFTTEAMTRY	2765
QY	1741	SAPPGDPPKPEYDLELITSCSNVSVAHDSAGKRYVYLTRDPTTLPARAAWETARHTPVN	1800
Db	2766	SAPPGDPPKPEYDLELITSCSNVSVAHDSAGKRYVYLTRDPTTLPARAAWETARHTPVN	2825
QY	1801	SWLGNIIWYATTLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIQL	1860
Db	2826	SWLGNIIWYATTLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPOIQL	2885
QY	1861	HGLSAFSLHSYSPGEINRVASCLRLKGLVPPFLRVWRHRRARSVRARLLSQGGRAAATCGKYL	1920
Db	2886	HGLSAFSLHSYSPGEINRVASCLRLKGLVPPFLRVWRHRRARSVRARLLSQGGRAAATCGKYL	2945
QY	1921	NWAVRTKULKPTIPAASQLDLSWTFVAGYSGGDIYHSLSRAPRPFMCLLLLSVGVGII	1980
Db	2946	NWAVRTKULKPTIPAASQLDLSWTFVAGYSGGDIYHSLSRAPRPFMCLLLLSVGVGII	3005

QY	1981	LLPNR	1985
Db	3006	LLPNR	3010
RESULT 9			
US-09-539-601-18			
; Sequence 18, Application US/09539601C			
; Patent No. 6630343			
; GENERAL INFORMATION:			
; APPLICANT: Bartenschlager, Ralf FW			
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System			
; FILE REFERENCE: all sequences			
; CURRENT APPLICATION NUMBER: US/09/539,601C			
; CURRENT FILING DATE: 2001-08-30			
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY			
; EARLIER FILING DATE: 1999-04-03			
; NUMBER OF SEQ ID NOS: 51			
; SOFTWARE: PatentIn Ver. 2.1			
; SEQ ID NO 18			
; LENGTH: 1985			
; TYPE: PR1			
; ORGANISM: Hepatitis C virus			
US-09-539-601-18			
Query Match 99.6%; Score 10425; DB 4; Length 1985;			
Best Local Similarity 99.6%; Pred. No. 0;			
Matches 1977; Conservative 4; Mismatches 4; Indels 0; Gaps 0;			
QY	1	MAPITAYSQQTGRLGCIITSLGRDRNQVEGVQVVSATQSFATCVNGVCHWTVYHGA	60
Db	1	MAPITAYSQQTGRLGCIITSLGRDRNQVEGVQVVSATQSFATCVNGVCHWTVYHGA	60
QY	61	GSKTLAGPKPIQMTYNDODLVGWCAPPKASLTCTCGSSDLYLVTRHADVIPVRRR	120
Db	61	GSKTLAGPKPIQMTYNDODLVGWCAPPKASLTCTCGSSDLYLVTRHADVIPVRRR	120
QY	121	GDGRGSLSPRPVSYLKGSSGGLPCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180
Db	121	GDGRGSLSPRPVSYLKGSSGGLPCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180
QY	181	RSVPFTDNSSPPAVPQTFQVAHLHAPTGSGSKSTKVPAAYAAQGYKVLVLPNSVAATLFGF	240
Db	181	RSVPFTDNSSPPAVPQTFQVAHLHAPTGSGSKSTKVPAAYAAQGYKVLVLPNSVAATLFGF	240
QY	241	AYMSKAHGIQPNRTGVTITTCGAPITYSTYKFLADGGCGGGAYDIIICDECHSTDSST	300
Db	241	AYMSKAHGIQPNRTGVTITTCGAPITYSTYKFLADGGCGGGAYDIIICDECHSTDSST	300
QY	301	ILGIGTVLDQAEATAGARLVWLATATPPGSVTVPHNIEEVALSSTGIPFYGKAIPETI	360
Db	301	ILGIGTVLDQAEATAGARLVWLATATPPGSVTVPHNIEEVALSSTGIPFYGKAIPETI	360
QY	361	KGGSHLIFCHSKKFCDELAIAKLSGLGLNNAVAYRGLDVSIVPTSGDVIVVATDALMTGFT	420
Db	361	KGGSHLIFCHSKKFCDELAIAKLSGLGLNNAVAYRGLDVSIVPTSGDVIVVATDALMTGFT	420
QY	421	GDPSVIDCNCTVQTVDPSLDPTFTTETTTVPQDAVRSORRGRTGGRMGIVRFVTPG	480
Db	421	GDPSVIDCNCTVQTVDPSLDPTFTTETTTVPQDAVRSORRGRTGGRMGIVRFVTPG	480
QY	481	ERPSSGMPDSSVLCFCYDAGCAWYBELTPAETSVRLRAYLNTPLGVPQDHLFWFSVFTGL	540
Db	481	ERPSSGMPDSSVLCFCYDAGCAWYBELTPAETSVRLRAYLNTPLGVPQDHLFWFSVFTGL	540
QY	541	THIDAHFLSQTKQAGDNFPLYVAYQATVCABAQAPPPSDOMKCLIRLKPTELHGPTPL	600
Db	541	THIDAHFLSQTKQAGDNFPLYVAYQATVCABAQAPPPSDOMKCLIRLKPTELHGPTPL	600
QY	601	YRLGAVQNEVTTTHPIKTIYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSSVIVGR	660
Db	601	YRLGAVQNEVTTTHPIKTIYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSSVIVGR	660

```
QY 661 IILSGKPAIPDREVLYRFEDEMEECASHLPYLEQGMQLAEQPKAKIQLQATATKQAEA 720
Db 661 IILSGKPAIPDREVLYRFEDEMEECASHLPYLEQGMQLAEQPKAKIQLQATATKQAEA 720
QY 721 AAPVVESKWTLEAFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTOH 780
Db 721 AAPVVESKWTLEAFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTOH 780
QY 781 TILFNILGGMVAQAOLAPPAASAFAVAGAGIAGAAVSGIGLKVLVDILAGYGAGVAGALVA 840
Db 781 TILFNILGGMVAQAOLAPPAASAFAVAGAGIAGAAVSGIGLKVLVDILAGYGAGVAGALVA 840
QY 841 FKVMGEMPTSDLVNLLPAILSPGALVVGWVCAAALRRHVGPGEAGVOMNRLIAFASR 900
Db 841 FKVMGEMPTSDLVNLLPAILSPGALVVGWVCAAALRRHVGPGEAGVOMNRLIAFASR 900
QY 901 GNVSPTHVPSDDAAARVTQILSSITITQLKRLHQWINEDECSTPCSGSMLRDVMDWIC 960
Db 901 GNVSPTHVPSDDAAARVTQILSSITITQLKRLHQWINEDECSTPCSGSMLRDVMDWIC 960
QY 961 TVLTDFKWLQSKLLPRLPGVFFFCORGYKGVWRGDGIMQITTCPGAQITGHVKNXGMR 1020
Db 961 TVLTDFKWLQSKLLPRLPGVFFFCORGYKGVWRGDGIMQITTCPGAQITGHVKNXGMR 1020
QY 1021 IVGPRTCSTNTHGTTPINAYTTGCTPSPAPNYSRALMEVAEEYVEVTRVGDHVVTCM 1080
Db 1021 IVGPRTCSTNTHGTTPINAYTTGCTPSPAPNYSRALMEVAEEYVEVTRVGDHVVTCM 1080
QY 1081 TTDNVKCPQVPAPEFFTEVDGVRHRYAPACKPLREEVTVLVLNQLVYLSQPCPBE 1140
Db 1081 TTDNVKCPQVPAPEFFTEVDGVRHRYAPACKPLREEVTVLVLNQLVYLSQPCPBE 1140
QY 1141 PDVAVLTSMLTDPSSHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTRHDSPAD 1200
Db 1141 PDVAVLTSMLTDPSSHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTRHDSPAD 1200
QY 1201 LIEANLLWRQMGNGNITRVESENKVVILDSFPLQAEDEREVSPAEILRRSRKFFPRAM 1260
Db 1201 LIEANLLWRQMGNGNITRVESENKVVILDSFPLQAEDEREVSPAEILRRSRKFFPRAM 1260
QY 1261 PIWARPDPNPLLESWKDDPVVPPVHGCPLPPAKAPPIPPRRKRTTVLSESTVSSALA 1320
Db 1261 PIWARPDPNPLLESWKDDPVVPPVHGCPLPPAKAPPIPPRRKRTTVLSESTVSSALA 1320
QY 1321 ELATKTFGSSSESSAVDSGTATASPDQSDGDAGSDVESYSPMPLEGEPPDLSGWS 1380
Db 1321 ELATKTFGSSSESSAVDSGTATASPDQSDGDAGSDVESYSPMPLEGEPPDLSGWS 1380
QY 1381 STVSEASEDEVVCCSMTYWTGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS 1440
Db 1381 STVSEASEDEVVCCSMTYWTGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS 1440
QY 1441 LRQKVTFDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEBACKLTPPHSARSFGYGA 1500
Db 1441 LRQKVTFDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEBACKLTPPHSARSFGYGA 1500
QY 1501 DVRNLSSKANVNHISVKNKLLLEDTEPTDITIMAKNEVFCVQPKGGRKPARLLVFPD 1560
Db 1501 DVRNLSSKANVNHISVKNKLLLEDTEPTDITIMAKNEVFCVQPKGGRKPARLLVFPD 1560
QY 1561 VRVCEKMALXDVNSTLPQAVMGSSYGFQSPQORVEFLVNAWAKKCPMGFPAYDTRCFDS 1620
Db 1561 VRVCEKMALXDVNSTLPQAVMGSSYGFQSPQORVEFLVNAWAKKCPMGFPAYDTRCFDS 1620
QY 1621 TVTENDIRVEESIYQCCDLAPAPQAIRSLTERLYIGPLTNSKGQNGYRRCRASGVLT 1680
Db 1621 TVTENDIRVEESIYQCCDLAPAPQAIRSLTERLYIGPLTNSKGQNGYRRCRASGVLT 1680
QY 1681 TSCGNTLTCVYKAAACRAAKLOCTMLVCGDDILVVICESAGTOEDEASIRAFTEAMTRY 1740
Db 1681 TSCGNTLTCVYKAAACRAAKLOCTMLVCGDDILVVICESAGTOEDEASIRAFTEAMTRY 1740
QY 1741 SAPPGDPPKPEYDLELITSCSSNVSAHDASGKRVYLTDRPTTTPPLARAWEATARHTPVN 1800
```

```
Db 1741 SAPPGDPPKPEYDLELITSCSSNVSAHDASGKRVYLTDRPTTTPPLARAWEATARHTPVN 1800
QY 1801 SWLGNLIIMYAPTILWARMILMTHFFSILLAQEOLEKALDCQIYGCACYSIEPLDLPQI 1860
Db 1801 SWLGNLIIMYAPTILWARMILMTHFFSILLAQEOLEKALDCQIYGCACYSIEPLDLPQI 1860
QY 1861 HGLSAFSLSHSYSPGEINRVASCLRLKLGVPPLRVHRHARSVRARLLSQGGRATATCKYLF 1920
Db 1861 HGLSAFSLSHSYSPGEINRVASCLRLKLGVPPLRVHRHARSVRARLLSQGGRATATCKYLF 1920
QY 1921 NNAVTRKLTPIPAASQDLSSWFWAGVSGGDIYHSLSRARPRMFWMCLLLSVGVGIY 1980
Db 1921 NNAVTRKLTPIPAASQDLSSWFWAGVSGGDIYHSLSRARPRMFWMCLLLSVGVGIY 1980
QY 1981 LLPNR 1985
Db 1981 LLPNR 1985

RESULT 10
US-09-539-601-21
; Sequence 21, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartenschlager, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-21

Query Match 99.6%; Score 10422; DB 4; Length 3010;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1976; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MAPITAYSQQTRGLLGLIITSLTGRDRNOVEGEVQVWSTATQSFATCVCNGVCTVYHGA 60
Db 1026 LAPITAYSQQTRGLLGLIITSLTGRDRNOVEGEVQVWSTATQSFATCVCNGVCTVYHGA 1085
QY 61 GSKTLAGPKGPIITQMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLYLVRHADVIPVRRR 120
Db 1086 GSKTLAGPKGPIITQMYTNVDQDLVGMQAPPGARSLTPTCTGSSDLYLVRHADVIPVRRR 1145
QY 121 GDSRGSLLSRPVSYLKSGSGPPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 180
Db 1146 GDSRGSLLSRPVSYLKSGSGPPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 1205
QY 181 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSGKSTKPAAYAAQGYKVLVLPNSVAATLGRG 240
Db 1206 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSGKSTKPAAYAAQGYKVLVLPNSVAATLGRG 1265
QY 241 AYMSKAHIGDPINRITGVRTITTCAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 300
Db 1266 AYMSKAHIGDPINRITGVRTITTCAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 1325
QY 301 ILGIGTVLDOAETAGARLVVLATATPGSVTVPHPNIEEVALSGTGPYKAIPETI 360
Db 1326 ILGIGTVLDOAETAGARLVVLATATPGSVTVPHPNIEEVALSGTGPYKAIPETI 1385
QY 361 KGRHLLIFCHSKKKKDELAALKSLGLNAVYVYRGLDVSVIPTSGDVIVVATATAMTGFT 420
Db 1386 KGRHLLIFCHSKKKKDELAALKSLGLNAVYVYRGLDVSVIPTSGDVIVVATATAMTGFT 1445
QY 421 GDFDSVIDNCVTQTVDFSLDPTFTTITTTVPQDAVSRSQRRTGRGRMGIYRFVTPG 480
```

```
1446 GDFSDVDCNTCVTQVDFSLDFTFFIEITVTPQDAVSRSQRGRGRGMYIRVFTPG 1505
1481 ERPSGMFSDSVLCECYDAGCAWYELTPAETSVELRAYLNTPLPVCQDHLBFWESVFTGL 540
1506 ERPSGMFSDSVLCECYDAGCAWYELTPAETSVELRAYLNTPLPVCQDHLBFWESVFTGL 1565
541 THIDAHLFSGTKOAGDNFPYLVAIQATVCARAQAPPSWDQWKCLIRLKPHTLHGPTLL 600
1566 THIDAHLFSGTKOAGDNFPYLVAIQATVCARAQAPPSWDQWKCLIRLKPHTLHGPTLL 1625
601 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVVTGTVLWVGVLAAALAAAYCLTTSVVIIVGR 660
1626 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVVTGTVLWVGVLAAALAAAYCLTTSVVIIVGR 1685
661 IILSGKPAIIPDREVLRYREFDEMEBCASHLPYIEQGMQLAEQFKQKAIQLQTATQAAEA 720
1686 IILSGKPAIIPDREVLRYREFDEMEBCASHLPYIEQGMQLAEQFKQKAIQLQTATQAAEA 1745
721 AAPVWESKWTLEAFNAKHMWNTISGTYLAGLSTLPGNPAIASLMAFTASITSPLTQH 780
1746 AAPVWESKWTLEAFNAKHMWNTISGTYLAGLSTLPGNPAIASLMAFTASITSPLTQH 1805
781 TLLFNILGWWAAQALAPPSAASAFVAGIAGAAVGSIGLQKVLVDIILAGYAGVAGALVA 840
1806 TLLFNILGWWAAQALAPPSAASAFVAGIAGAAVGSIGLQKVLVDIILAGYAGVAGALVA 1865
841 FKVMGMPSTEDLVNLLPAILSPGALVGVVCAAILRHVGPGEAGVQWNNRLIAPASR 900
1866 FKVMGMPSTEDLVNLLPAILSPGALVGVVCAAILRHVGPGEAGVQWNNRLIAPASR 1925
901 GNVSPHYVPESDAAARVQIILSSLTITQLKELHOMINEDCSTPCSGWLRDWDWIC 960
1926 GNVSPHYVPESDAAARVQIILSSLTITQLKELHOMINEDCSTPCSGWLRDWDWIC 1985
961 TVLTDFTKTLQSKLLPRLPGVPFFSCQYKGVWRGDMQTTCCPGAQITGHVKNQSMR 1020
1986 TVLTDFTKTLQSKLLPRLPGVPFFSCQYKGVWRGDMQTTCCPGAQITGHVKNQSMR 2045
1021 IVPRTCSNTHWGFPPINAYTGTCTSPAPNYGRALWRVAABEYVEVTRVGDHYHTGM 1080
2046 IVPRTCSNTHWGFPPINAYTGTCTSPAPNYGRALWRVAABEYVEVTRVGDHYHTGM 2105
1081 TTDNVKCPQVPAPEFTEVDGVHLRYAPACKPLAREEVTFLVGLNQYLWGSOLPCPE 1140
2106 TTDNVKCPQVPAPEFTEVDGVHLRYAPACKPLAREEVTFLVGLNQYLWGSOLPCPE 2165
1141 PDVAVLTSMULTDPSHITAETAKRRLARGSPPSLASSASQSLAPSLKATCTTRHDSPAD 1200
2166 PDVAVLTSMULTDPSHITAETAKRRLARGSPPSLASSASQSLAPSLKATCTTRHDSPAD 2225
1201 LIEANLLWRQMGNGNI TRVESENKVILDSPEPIQAEDEDEREVSVPABILRRSRKFFPRAM 1260
2226 LIEANLLWRQMGNGNI TRVESENKVILDSPEPIQAEDEDEREVSVPABILRRSRKFFPRAM 2285
1261 PIWARPDPNPLLESWKPDPVVPVHGCPLPPAKAPPIPPRRKRTVVLSESVSSALA 1320
2286 PIWARPDPNPLLESWKPDPVVPVHGCPLPPAKAPPIPPRRKRTVVLSESVSSALA 2345
1321 ELATKTFGSSSESSAVDSGTATASPDQSDDGAGSDVESYMSMPLEGEPPDLSGSGW 1380
2346 ELATKTFGSSSESSAVDSGTATASPDQSDDGAGSDVESYMSMPLEGEPPDLSGSGW 2405
1381 STVSEASEDVVCCSMSTWTGALITPCAABETKLPINALSNSLLRHNLVYATTSRAS 1440
2406 STVSEASEDVVCCSMSTWTGALITPCAABETKLPINALSNSLLRHNLVYATTSRAS 2465
1441 LRQKVTEDRLQVLDHVRDVLKENKAKASTVKAKLLSVEACKLTPPHSARSKFYGAK 1500
2466 LRQKVTEDRLQVLDHVRDVLKENKAKASTVKAKLLSVEACKLTPPHSARSKFYGAK 2525
1501 DVNRLSSKAVNHIRSVMKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLG 1560

2526 DVNRLSSKAVNHIRSVMKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIIVFPDLG 2585
1561 VRVCEKMAIYDVYSTLTPQAVMGSSYGFQYSPGQVFEFLVNAWKAKKCPMGFAYDTRCFDS 1620
2586 VRVCEKMAIYDVYSTLTPQAVMGSSYGFQYSPGQVFEFLVNAWKAKKCPMGFAYDTRCFDS 2645
1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRRCRAGVLT 1680
2646 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRRCRAGVLT 2705
1681 TSCGNTLTLCYLKAAACRAAKLOCTMLVCGDDLVI CESAGTQDEDEASLRAFTEAMTRY 1740
2706 TSCGNTLTLCYLKAAACRAAKLOCTMLVCGDDLVI CESAGTQDEDEASLRAFTEAMTRY 2765
1741 SAPPGPPKPEYDLELITSCSSNVSAHDASGRVYVLTDPDPTPLARAAWETARHTPVN 1800
2766 SAPPGPPKPEYDLELITSCSSNVSAHDASGRVYVLTDPDPTPLARAAWETARHTPVN 2825
1801 SWLGNTIMVAPTILWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORL 1860
2826 SWLGNTIMVAPTILWARMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORL 2885
1861 HGLSAPLSHSYSGEINRVASCLKLGVPPLRVWRHRSVRARLLSQGGRAATCGKYL 1920
2886 HGLSAPLSHSYSGEINRVASCLKLGVPPLRVWRHRSVRARLLSQGGRAATCGKYL 2945
1921 NWAVRTKLTPTIPAAASQDLSSWVAGYSGGDIYHLSRARPFWFMCCLLLSVGVGIY 1980
2946 NWAVRTKLTPTIPAAASQDLSSWVAGYSGGDIYHLSRARPFWFMCCLLLSVGVGIY 3005
1981 LLPNR 1985
3006 LLPNR 3010

RESULT 11
US-09-539-601-30
; Sequence 30, Application US/09539601C
; Patent No. 6630343
; GENERAL INFORMATION:
; APPLICANT: Bartschlagel, Ralf FW
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System
; FILE REFERENCE: all sequences
; CURRENT APPLICATION NUMBER: US/09/539,601C
; CURRENT FILING DATE: 2001-08-30
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY
; EARLIER FILING DATE: 1999-04-03
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 30
; LENGTH: 1985
; TYPE: PRT
; ORGANISM: Hepatitis C virus
US-09-539-601-30

Query Match 99.5%; Score 10412; DB 4; Length 1985;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1974; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 MAPITAYSQTRGLLGCITITSLTGRDRNQVEGEVQVSTATQSFATCCVGCMTVYHGA 60
Db 1 MAPITAYSQTRGLLGCITITSLTGRDRNQVEGEVQVSTATQSFATCCVGCMTVYHGA 60
QY 61 GSKTLAGPKGITQMYTNVDQDLVGWQAPPCARSITPTCTCGSSDLVLTVRHADVIPVRRR 120
Db 61 GSKTLAGPKGITQMYTNVDQDLVGWQAPPCARSITPTCTCGSSDLVLTVRHADVIPVRRR 120
QY 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPGSHAVGIPFRAAVCTRGVAKAVDFVPVESMETTM 180
Db 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPGSHAVGIPFRAAVCTRGVAKAVDFVPVESMETTM 180
QY 181 RSPVFTDSSPPAVPQTFQVAHLHAPTSGKSTKVPAAVAAQGYKVLVNPVAAATLFG 240
Db 181 RSPVFTDSSPPAVPQTFQVAHLHAPTSGKSTKVPAAVAAQGYKVLVNPVAAATLFG 240
```

181 RSPVFTDNGSPAPQVTFQVAHLHPTGSGKSTKVEAAVAGQGYKVLVNLNPSVAATLFGF 240  
241 AYMSKAHGIDPNIRITGVRITTCGAPITTSYTGKFLADGGCGSGAYDIIICDECHSDSTT 300  
241 AYMSKAHGIDPNIRIGVRITTCGAPITTSYTGKFLADGGCGSGAYDIIICDECHSDSTT 300  
301 ILGIGTVLDOAETAGARLVVLATATPPGVTVPHPNIEEVALSSCTGEIPFYGKAIPETI 360  
301 ILGIGTVLDOAETAGARLVVLATATPPGVTVPHPNIEEVALSSCTGEIPFYGKAIPETI 360  
361 KGRHLIFCHSKKCDLAKSLGSLGNNAVAYRGLDVSVIPTS GDVIVVATDALMTGFT 420  
361 KGRHLIFCHSKKCDLAKSLGSLGNNAVAYRGLDVSVIPTS GDVIVVATDALMTGFT 420  
421 GDFSDVIDCNTCTQVDFSLDPTFTIETTTVPQDAVSQRGRGRCGRMGIRFYVTPG 480  
421 GDFSDVIDCNTCTQVDFSLDPTFTIETTTVPQDAVSQRGRGRCGRMGIRFYVTPG 480  
481 BRPSGMFSSVLCYDAGCAWYELTPAETSVRLRAYLNTGTLPGVQDHLFEWESVFTGL 540  
481 BRPSGMFSSVLCYDAGCAWYELTPAETSVRLRAYLNTGTLPGVQDHLFEWESVFTGL 540  
541 THIDAHFUSQTKOAGDNPPYLVAQTVCARAQAPPPSWDQWKCLIRLKP TLHGFTPLL 600  
541 THIDAHFUSQTKOAGDNPPYLVAQTVCARAQAPPPSWDQWKCLIRLKP TLHGFTPLL 600  
601 YRLGAVQNEVTTTHPIKYIMACMSADLEVTSTWLVGGVLAALAAAYCLITGSSVVIYGR 660  
601 YRLGAVQNEVTTTHPIKYIMACMSADLEVTSTWLVGGVLAALAAAYCLITGSSVVIYGR 660  
661 IILSKPAIIPDREVLRYEFDEMEBCASHLPYIEQGMQLAEQFKQKAGLLQATKQABE 720  
661 IILSKPAIIPDREVLRYEFDEMEBCASHLPYIEQGMQLAEQFKQKAGLLQATKQABE 720  
721 AAPVVEKWTLEAPWAKHWNFTSGIQLAGSLTPGNDPAIASLMAFTASITSPLTQH 780  
721 AAPVVEKWTLEAPWAKHWNFTSGIQLAGSLTPGNDPAIASLMAFTASITSPLTQH 780  
781 TLLFNILGGWAAOLAPPSAASAFVAGIAGAAVGSIGLKVLDVILAGYGAGVAGALVA 840  
781 TLLFNILGGWAAOLAPPSAASAFVAGIAGAAVGSIGLKVLDVILAGYGAGVAGALVA 840  
841 FKVMSGEMPSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVQVWNNRLIAFASR 900  
841 FKVMSGELPSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVQVWNNRLIAFASR 900  
901 GNHVSPTHYVPESDAAARVTOILSLTITQLKRLHWINEDCSTPCGSLRDLVDWVIC 960  
901 GNHVSPTHYVPESDAAARVTOILSLTITQLKRLHWINEDCSTPCGSLRDLVDWVIC 960  
961 TVLTDFTLWQSKLLPRLPGVPFFSCORGKYGVRGDMOTTCPCGAQIITGHVKNQSMR 1020  
961 TVLTDFTLWQSKLLPRLPGVPFFSCORGKYGVRGDMOTTCPCGAQIITGHVKNQSMR 1020  
1021 IVGPRTCSNTWHGTFPINAIVTGPCTSPAPNYGRALWRVAEEYVEVTRVGDPHYVTGM 1080  
1021 IVGPRTCSNTWHGTFPINAIVTGPCTSPAPNYGRALWRVAEEYVEVTRVGDPHYVTGM 1080  
1081 TTNVVKCPQVPAPEFTEVDGVRHLRYAPACKPLLEEVTVLVLNQYLVGSLQPCPE 1140  
1081 TTNVVKCPQVPAPEFTEVDGVRHLRYAPACKPLLEEVTVLVLNQYLVGSLQPCPE 1140  
1141 PDVAVLTSMLTDSHITAETAKRLARGSPSSASSASOLSAPLSKATCTTRHDSPAD 1200  
1141 PDVAVLTSMLTDSHITAETAKRLARGSPSSASSASOLSAPLSKATCTTRHDSPAD 1200  
1201 LIEANLWQEMGNIITRVESENKVVILDSFEPLQAEDEDEREVSVPAAEILRRKRKPPRAM 1260  
1201 LIEANLWQEMGNIITRVESENKVVILDSFEPLQAEDEDEREVSVPAAEILRRKRKPPRAM 1260  
1261 PIWAPDYNPPLLESWKDPDYPVPPVHGCPLPPAKAPPPIPPRRKRTVILSESTVSALA 1320  
1261 PIWAPDYNPPLLESWKDPDYPVPPVHGCPLPPAKAPPPIPPRRKRTVILSESTVSALA 1320

1321 ELATKTFGSSSANDSGTATSPDQPSDDGAGSDVESYSSMPLEGBEGDPLDGGSW 1380  
1321 ELATKTFGSSSANDSGTATSPDQPSDDGAGSDVESYSSMPLEGBEGDPLDGGSW 1380  
1381 STVSEASESDVCCSMSTYTWTGALITPCAETKLPINALSNLLRHHNLVYATTSRSAS 1440  
1381 STVSEASESDVCCSMSTYTWTGALITPCAETKLPINALSNLLRHHNLVYATTSRSAS 1440  
1441 LRQKKVTFRLOVLDDHVDLKEKAKASTVKAKLLSVEEACKLTPPHSARSKFYGA 1500  
1441 LRQKKVTFRLOVLDDHVDLKEKAKASTVKAKLLSVEEACKLTPPHSARSKFYGA 1500  
1501 DVRLNSSKAVNHRSVWKLLEDTEPTDITTIMAKNEVEFCVQPEKGRKPARLIVFPDLG 1560  
1501 DVRLNSSKAVNHRSVWKLLEDTEPTDITTIMAKNEVEFCVQPEKGRKPARLIVFPDLG 1560  
1561 VRVCEKMALYDVVSTLPQAVMSSYGFQYSPQORVEFLVNAWKKCPMGFAYDTRCFDS 1620  
1561 VRVCEKMALYDVVSTLPQAVMSSYGFQYSPQORVEFLVNAWKKCPMGFAYDTRCFDS 1620  
1621 TVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKGNCGYRECRASGVLT 1680  
1621 TVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKGNCGYRECRASGVLT 1680  
1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVI CESAGTQDEASLRAFTAMTRY 1740  
1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVI CESAGTQDEASLRAFTAMTRY 1740  
1741 SAPGDPKPKPYDLELITSCSNVSVAHDASGRVYVYTRDPTTPPLARAAMTARHTPVN 1800  
1741 SAPGDPKPKPYDLELITSCSNVSVAHDASGRVYVYTRDPTTPPLARAAMTARHTPVN 1800  
1801 SWLGNIMYAPTLWARMLMTHFTSILLAOBOLKALDCQIYGACYSIEPLDLPOLIORL 1860  
1801 SWLGNIMYAPTLWARMLMTHFTSILLAOBOLKALDCQIYGACYSIEPLDLPOLIORL 1860  
1861 HGLSAPLSHVSPEINRVSCLRKLGVPLRVHRARSVRARLLSOGGRAATCGKYL 1920  
1861 HGLSAPLSHVSPEINRVSCLRKLGVPLRVHRARSVRARLLSOGGRAATCGKYL 1920  
1921 NNAVTKLTPIPAASQLDLSWFFVAGYSGDIIYHLSRARPWFWMCLLLSVGVGIY 1980  
1921 NNAVTKLTPIPAASQLDLSWFFVAGYSGDIIYHLSRARPWFWMCLLLSVGVGIY 1980  
1981 LLPNR 1985  
1981 LLPNR 1985

## RESULT 12

US-09-539-601-33  
; Sequence 33, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 33  
; LENGTH: 3010  
; TYPE: PRN  
; ORGANISM: Hepatitis C virus  
US-09-539-601-33

Query Match 99.5%; Score 10409; DB 4; Length 3010;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 1973; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY	1	MAPITAYSQOTRGLLCIIITSLTGRDRNQVEGVQVSTATOSFLATCNGVCWTYHGA	60	QY	1081	TTDNVCKPCQVPAPPEFFTEVDGVRLHRYAPACKPLLRREVTFLVGLNQVLVGSQLPCEPE	1140
Db	1026	LAPITAYSQOTRGLLCIIITSLTGRDRNQVEGVQVSTATOSFLATCNGVCWTYHGA	1085	Db	2106	TTDNVCKPCQVPAPPEFFTEVDGVRLHRYAPACKPLLRREVTFLVGLNQVLVGSQLPCEPE	2165
QY	61	GSKTLAGPGPITOMYTNVDODLVGQAPGARSLTPTCGSSDLVLRHADVIPRRR	120	QY	1141	PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLASSASQLSAPSILKATCTTHRHSDDAD	1200
Db	1086	GSKTLAGPGPITOMYTNVDODLVGQAPGARSLTPTCGSSDLVLRHADVIPRRR	1145	Db	2166	PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLASSASQLSAPSILKATCTTHRHSDDAD	2225
QY	121	GDSRGSLSPRPVSYLKGSSGGPLCPGSHAVGIFRAAVCTRGVAKAVDFPVVPMETTM	180	QY	1201	LIEANLLWQEMQGNITRVSEENKVILDSFEPLQAEEDEREVSVAEILRRRKRKPPRAM	1260
Db	1146	GDSRGSLSPRPVSYLKGSSGGPLCPGSHAVGIFRAAVCTRGVAKAVDFPVVPMETTM	1205	Db	2226	LIEANLLWQEMQGNITRVSEENKVILDSFEPLQAEEDEREVSVAEILRRRKRKPPRAM	2285
QY	181	RSVFDTNDSPPAVPOTFOVAHLHAPGSGKSTKVPAAYAAQGYKVLVLPNSVAATLFGF	240	QY	1261	PIWARDYNPPLLESWKDPDYVPPVHVGCPLPPAKAPPIPPRRKRTVVLSESTVSALA	1320
Db	1206	RSVFDTNDSPPAVPOTFOVAHLHAPGSGKSTKVPAAYAAQGYKVLVLPNSVAATLFGF	1265	Db	2286	PIWARDYNPPLLESWKDPDYVPPVHVGCPLPPAKAPPIPPRRKRTVVLSESTVSALA	2345
QY	241	AYMSKAGHDIPNIRIGVTRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST	300	QY	1321	ELATKTFGSSAVDSGTATASPDOPSDDDGAGSDVESYSSMPPLEGEGPDPLSDGSW	1380
Db	1266	AYMSKAGHDIPNIRIGVTRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDST	1325	Db	2346	ELATKTFGSSAVDSGTATASPDOPSDDDGAGSDVESYSSMPPLEGEGPDPLSDGSW	2405
QY	301	ILGIGTVLQDAETAGARLVLATATPPGSVTVPHPNIEEVALSGTGEBIPFYKAIPIETI	360	QY	1381	STVSEASEDDVCCMSYTWGTALITPCAABETKLPIINALSNLLRHNLVYATTGRSAS	1440
Db	1326	ILGIGTVLQDAETAGARLVLATATPPGSVTVPHPNIEEVALSGTGEBIPFYKAIPIETI	1385	Db	2406	STVSEASEDDVCCMSYTWGTALITPCAABETKLPIINALSNLLRHNLVYATTGRSAS	2465
QY	361	KGRHILFCHSKKKKDELAALSGLGNLNAVAYRGLDVSVIPTSGLDVIVVATDALMTGFT	420	QY	1441	LROKKTFTORLQVLDHRYDLVKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYGAK	1500
Db	1386	KGRHILFCHSKKKKDELAALSGLGNLNAVAYRGLDVSVIPTSGLDVIVVATDALMTGFT	1445	Db	2466	LROKKTFTORLQVLDHRYDLVKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYGAK	2525
QY	421	GDPSVLDNCTCVTOTWDFSLDPTFTTTPQDVSRRORRGTGRGMGIYRFVTPG	480	QY	1501	DVRNLSKAVNHITRSVWKDLLEDTETPIDTTIIVAKNEVFCVQPEKGRKPARLIVFPDLG	1560
Db	1446	GDPSVLDNCTCVTOTWDFSLDPTFTTTPQDVSRRORRGTGRGMGIYRFVTPG	1505	Db	2526	DVRNLSKAVNHITRSVWKDLLEDTETPIDTTIIVAKNEVFCVQPEKGRKPARLIVFPDLG	2585
QY	481	ERPSGMEFSSVLCBCEYDAGCAWYELTPTAETSVRLRAYLNTPLPVCQDHLFEWESVFTGL	540	QY	1561	VRCEKMAVDVYSTLPQAVMGSSYFQYSPGQVRFVFNWAKAKKCPMGFAYDTRCFDS	1620
Db	1506	ERPSGMEFSSVLCBCEYDAGCAWYELTPTAETSVRLRAYLNTPLPVCQDHLFEWESVFTGL	1565	Db	2586	VRCEKMAVDVYSTLPQAVMGSSYFQYSPGQVRFVFNWAKAKKCPMGFAYDTRCFDS	2645
QY	541	THIDAHFLSOTKQAGDNFPYLVAYQAVCARAQAPPSWDQMWKCLRLKPTLHGPTLL	600	QY	1621	TVTENDIRVBESIVQCCLDAPARQAIIRSLTERLYIGGPLTNSKGQNCGRRCRASGVLT	1680
Db	1566	THIDAHFLSOTKQAGDNFPYLVAYQAVCARAQAPPSWDQMWKCLRLKPTLHGPTLL	1625	Db	2646	TVTENDIRVBESIVQCCLDAPARQAIIRSLTERLYIGGPLTNSKGQNCGRRCRASGVLT	2705
QY	601	YRLGAVQNEVTTTHPIITKYINACMSADLEVVTSTWLVGGVLAALAYCLTGSWIVGR	660	QY	1681	TSCGNTUTCVLKAACRAAKLQDCTMLVCGDDLVVTCESAGTQDEASLRAFTTEAMTRY	1740
Db	1626	YRLGAVQNEVTTTHPIITKYINACMSADLEVVTSTWLVGGVLAALAYCLTGSWIVGR	1685	Db	2706	TSCGNTUTCVLKAACRAAKLQDCTMLVCGDDLVVTCESAGTQDEASLRAFTTEAMTRY	2765
QY	661	IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLABQFKQKAILGLQTATKQAE	720	QY	1741	SAPGDPKPEYDLELITSCSSNVSVVAHDASGKRVYVLTTRDPTTPLARAAMETARHTPVN	1800
Db	1686	IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLABQFKQKAILGLQTATKQAE	1745	Db	2766	SAPGDPKPEYDLELITSCSSNVSVVAHDASGKRVYVLTTRDPTTPLARAAMETARHTPVN	2825
QY	721	AAPVVEKWRTEAFWAKHWNFTSGIOYLAGLSTLPGNPAIASIMAFASITSPITTOH	780	QY	1801	SWLGNIIIMYAPTLWARMILMTHFESILLAQOEKALDQCIYGCYCIETPDLQIIOQL	1860
Db	1746	AAPVVEKWRTEAFWAKHWNFTSGIOYLAGLSTLPGNPAIASIMAFASITSPITTOH	1805	Db	2826	SWLGNIIIMYAPTLWARMILMTHFESILLAQOEKALDQCIYGCYCIETPDLQIIOQL	2885
QY	781	TLLFNILGGWAAQOLAPPSAASAFVAGIAGAVSGIGLKVLDVILAGYAGVAGALVA	840	QY	1861	HGLSAPSLHSYSGPEINRVASCLRKLGVPPLRVWHRARSVRARLLSOGGAAATCGKYL	1920
Db	1806	TLLFNILGGWAAQOLAPPSAASAFVAGIAGAVSGIGLKVLDVILAGYAGVAGALVA	1865	Db	2886	HGLSAPSLHSYSGPEINRVASCLRKLGVPPLRVWHRARSVRARLLSOGGAAATCGKYL	2945
QY	841	FKVMSGEMPTEDLVNLLPALISPGALVGVVCAAILRRHVPGEVAVQWNNRLIAFASR	900	QY	1921	NWAVRTKLTPIPAASOLDLSSWVAGYSGGDIYHSLSRAPRPFWMCLLLSVGVGIY	1980
Db	1866	FKVMSGEMPTEDLVNLLPALISPGALVGVVCAAILRRHVPGEVAVQWNNRLIAFASR	1925	Db	2946	NWAVRTKLTPIPAASOLDLSSWVAGYSGGDIYHSLSRAPRPFWMCLLLSVGVGIY	3005
QY	901	GNHVSPTHYVPSDAAARVTOILLSSLTITOLLKRLHOMINEDCSTPCSGMLRDVMDWIC	960	QY	1981	LLPNR 1985	
Db	1926	GNHVSPTHYVPSDAAARVTOILLSSLTITOLLKRLHOMINEDCSTPCSGMLRDVMDWIC	1985	Db	3006	LLPNR 3010	

RESULT 13  
US-08-952-981A-2  
; Sequence 2, Application US/08952981A  
; Patent No. 6383768  
; GENERAL INFORMATION:  
; APPLICANT: DE FRANCESCO, Raffaele  
; APPLICANT: TOMELI, Licia  
; APPLICANT: BEHRENS, Sven-Erik



; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE RNA-DEPENDENT RNA  
; TITLE OF INVENTION: POLYMERASE AND TERMINAL NUCLEOTIDYL TRANSFERASE  
; FILE REFERENCE: IT0002P  
; CURRENT APPLICATION NUMBER: US/08/952,981A  
; CURRENT FILING DATE: 1998-03-23  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2201  
; TYPE: PRT  
; ORGANISM: cDNA clone pCD (38-9.4)  
US-08-952-981A-2

Query Match		97.2%;	Score 10170;	DB 3;	Length 2201;	
Best Local Similarity		96.7%;	Pred. No. 0;			
Matches 1919;		Conservative 36;	Mismatches 30;	Indels 0;	Gaps 0;	
QY	1	MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGA	60			
DB	217	LAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGA	276			
QY	61	GSKTLAGKPIITOMYTNVDQVLQVQAPPGARSITPCTCGSSDLYLVTRHADVIPVRRR	120			
DB	277	GSKTLAGKPIITOMYTNVDQVLQVQAPPGARSITPCTCGSSDLYLVTRHADVIPVRRR	336			
QY	121	GDSRGLSPRPVSVLKGSSGGPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180			
DB	337	GDSRGLSPRPVSVLKGSSGGPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	396			
QY	181	RSPVFTDNSSPFAVQTFQVAHLHAPTGSKSTKVPAAVAGYKVLVNPVSAATLGF	240			
DB	397	RSPVFTDNSSPFAVQTFQVAHLHAPTGSKSTKVPAAVAGYKVLVNPVSAATLGF	456			
QY	241	AYMSKAHGDINIRGVTITITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDST	300			
DB	457	AYMSKAHGDINIRGVTITITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDST	516			
QY	301	ILGTVLDQAEAGARLVVATATPPGVSVPVPHNIEVALSSTGEIPFYKALPIETI	360			
DB	517	ILGTVLDQAEAGARLVVATATPPGVSVPVPHNIEVALSSTGEIPFYKALPIETI	576			
QY	361	KGSRHLIFCHSKKCKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVLVWATDALMTGFT	420			
DB	577	KGSRHLIFCHSKKCKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVLVWATDALMTGFT	636			
QY	421	GDFDSVIDCNTCVTQTFDLSLDTFTTITTPQDAVSRSQRRGTRGMRGIYRFVTPG	480			
DB	637	GDFDSVIDCNTCVTQTFDLSLDTFTTITTPQDAVSRSQRRGTRGMRGIYRFVTPG	696			
QY	481	ERPSGMFDSVILCECYDAGCAWYELTPAETSVRLRAYINTGLPVQCQDHLFWESVFTGL	540			
DB	697	ERPSGMFDSVILCECYDAGCAWYELTPAETSVRLRAYINTGLPVQCQDHLFWESVFTGL	756			
QY	541	THIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAPPSWDMKCLIRLKPTELHGPTPL	600			
DB	757	THIDAHFLSQTKQAGDNFPYLVAYQATVCARAQAPPSWDMKCLIRLKPTELHGPTPL	816			
QY	601	YRLGAVQNEVTHPIITKYIMACMSADLEVVTSTWVLVGGVLAALAYCLTTGSSWIVGR	660			
DB	817	YRLGAVQNEVTHPIITKYIMACMSADLEVVTSTWVLVGGVLAALAYCLTTGSSWIVGR	876			
QY	661	IILSGKPAIIPDRVLVREFDEMEECASHLPYIEQGMQLAEQFKQKALGLQATKQAEA	720			
DB	877	IILSGKPAIIPDRVLVREFDEMEECASHLPYIEQGMQLAEQFKQKALGLQATKQAEA	936			
QY	721	AAPVVEKWRTELEAFWAKHWNFTSIGIYLAGLSTLPCNPALASLMAFTASITSLTQH	780			
DB	937	AAPVVEKWRTELEAFWAKHWNFTSIGIYLAGLSTLPCNPALASLMAFTASITSLTQH	996			
QY	781	TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLGVLDIILAGYGAGVAGALVA	840			
DB	997	TLLFNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLGVLDIILAGYGAGVAGALVA	1056			

QY	841	FKVMGEMSTEDLVNLLPAILSPGALVVGVCVCAAILRRHVGPGEGAVQMMNRLIAFASR	900
DB	1057	FKVMGEMSTEDLVNLLPAILSPGALVVGVCVCAAILRRHVGPGEGAVQMMNRLIAFASR	1116
QY	901	GNHVSPTHVPESDAAARVTQILSSLTITQLLKRHLQWINEDECSTPCSGSLRDVWDWIC	960
DB	1117	GNHVSPTHVPESDAAARVTQILSSLTITQLLKRHLQWINEDECSTPCSGSLRDVWDWIC	1176
QY	961	TVLTDPKTWLQSKLLPRLPGVPFFESCORGKGVWBGDMQTTCCGAQITGHVKNQSMR	1020
DB	1177	TVLTDPKTWLQSKLLPRLPGVPFFESCORGKGVWBGDMQTTCCGAQITGHVKNQSMR	1236
QY	1021	IVGPRTCSNTWHGTFFPINAYTTGCTPSPAPNYSRALMRVAEEVEVTRVGDHFVVTGM	1080
DB	1237	IVGPRTCSNTWHGTFFPINAYTTGCTPSPAPNYSRALMRVAEEVEVTRVGDHFVVTGM	1296
QY	1081	TTDNVKKCPQVPAPEFFTEVDGRLHRYAPACKPLLRREVTPLVGLNOYLVGSQLPCEPE	1140
DB	1297	TTDNVKKCPQVPAPEFFTEVDGRLHRYAPACKPLLRREVTPLVGLNOYLVGSQLPCEPE	1356
QY	1141	PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLASSASQLSAPSLKATCTTHHVSPPAD	1200
DB	1357	PDVAVLTSMLTDPSSHITAEATKRRLARGSPPSLASSASQLSAPSLKATCTTHHVSPPAD	1416
QY	1201	LIEANLLWRQBMGNGNITRVESENKVILDSPEPLQAEEDEREVSVPAILRRSRKFFPRAM	1260
DB	1417	LIEANLLWRQBMGNGNITRVESENKVILDSPEPLQAEEDEREVSVPAILRRSRKFFPRAM	1476
QY	1261	PIWARPDPNPLLSWKKDPDYVPVHGGCPPLPAKAPPIPPRRKRTVVLSESTVSSALA	1320
DB	1477	PIWARPDPNPLLSWKKDPDYVPVHGGCPPLPAKAPPIPPRRKRTVVLSESTVSSALA	1536
QY	1321	ELATKTFGSSBSAVDSGTATASDPQSDGDSGDSVESYSSMPLEGEPCDPLSDGSM	1380
DB	1537	ELATKTFGSSBSAVDSGTATASDPQSDGDSGDSVESYSSMPLEGEPCDPLSDGSM	1596
QY	1381	STVSEASESDVCCSMSTYTGTALITPCAAETKLIPINALSNSLLRHHNLVYATTSRAS	1440
DB	1597	STVSEASESDVCCSMSTYTGTALITPCAAETKLIPINALSNSLLRHHNLVYATTSRAS	1656
QY	1441	LROKVTDFRLQVLDHRYDLKEMKAKASTVKAALLSVEPACKLTPHSAKSKFGYGA	1500
DB	1657	LROKVTDFRLQVLDHRYDLKEMKAKASTVKAALLSVEPACKLTPHSAKSKFGYGA	1716
QY	1501	DVRNLSSKANVHJRSVMKDLLEDTEPTDTIMAKNEVFCVQPEKGGKPARLIVFPDLG	1560
DB	1717	DVRNLSSKANVHJRSVMKDLLEDTEPTDTIMAKNEVFCVQPEKGGKPARLIVFPDLG	1776
QY	1561	VRVCEKVALYDVVSTLTPQAVMGSSYGFQSPGQRFVEFLVNAWAKKCPMGPAYDTRCFDS	1620
DB	1777	VRVCEKVALYDVVSTLTPQAVMGSSYGFQSPGQRFVEFLVNAWAKKCPMGPAYDTRCFDS	1836
QY	1621	TVTENDIRVESIYQCCDLAPAPARQATRSUTERLYIGGPTLNSKGQCGVRRCRASGVL	1680
DB	1837	TVTENDIRVESIYQCCDLAPAPARQATRSUTERLYIGGPTLNSKGQCGVRRCRASGVL	1896
QY	1681	TSCGNTLTCVLKAAACRAAKLQDCTMLVCGDDLVLICESAGTQDEASLRAFTTEAMTRY	1740
DB	1897	TSCGNTLTCVLKAAACRAAKLQDCTMLVCGDDLVLICESAGTQDEASLRAFTTEAMTRY	1956
QY	1741	SAPPGPPPEYDLELITSCSNVSVVAHDASGKRVYVLTDRDPTTPLARAWEATARTHPVN	1800
DB	1957	SAPPGPPPEYDLELITSCSNVSVVAHDASGKRVYVLTDRDPTTPLARAWEATARTHPVN	2016
QY	1801	SWLGNIMVAPTILWARMILMTHFFSILLAQEULEKALDQCIYGCYSIEPLDLPQIQR	1860
DB	2017	SWLGNIMVAPTILWARMILMTHFFSILLAQEULEKALDQCIYGCYSIEPLDLPQIQR	2076
QY	1861	HGLSAFSLHSYSGEINRVASCLRLKGVPLRVHRARSVRARLLSQGGAATCGKYL	1920
DB	2077	HGLSAFSLHSYSGEINRVASCLRLKGVPLRVHRARSVRARLLSQGGAATCGKYL	2136

QY 1921 NNAVTKLKLTPIPASQDLSSWFWAGYSGGDIYHSLSRPRPFWMCCLLLSVGVGIY 1980  
 Db 2137 NNAVTKLKLTPIPASRLDLSGFWAGYSGGDIYHSLSRPRPFWMCCLLLSVGVGIY 2196  
 QY 1981 LLPNR 1985  
 Db 2197 LLPNR 2201

RESULT 14  
 US-08-324-977-36  
 ; Sequence 36, Application US/08324977  
 ; Patent No. 5747339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: OKAYAMA, Hiroto  
 ; APPLICANT: FURE, Isao  
 ; APPLICANT: MORI, Chisato  
 ; APPLICANT: TAKAMIZAWA, Akahisa  
 ; APPLICANT: YOSHIDA, Iwao  
 ; TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
 ; TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
 ; NUMBER OF SEQUENCES: 50  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Armstrong, Westerman, Hattori, McLeLeland &  
 ; ADDRESSEE: Naughton  
 ; STREET: 1725 K St. N.W. Suite 1000  
 ; CITY: Washington  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20006  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
 ; SOFTWARE: ASCII  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/324,977  
 ; FILING DATE: 18-OCT-1994  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 2-167466  
 ; FILING DATE: 25-JUN-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 2-230921  
 ; FILING DATE: 31-AUG-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: JP 2-305605  
 ; FILING DATE: 09-NOV-1990  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 08/099,706  
 ; FILING DATE: 30-JUL-1993  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/769,996  
 ; FILING DATE: 02-OCT-1991  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 07/635,451  
 ; FILING DATE: 28-DEC-1990  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Stevens-Smith, Theresa M.  
 ; REGISTRATION NUMBER: 36,281  
 ; REFERENCE/DOCKET NUMBER: 900703D  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 659-2930  
 ; TELEFAX: (202) 887-0357  
 ; TELEX: 440142  
 ; INFORMATION FOR SEQ ID NO: 36:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 2621 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-08-324-977-36

Query Match 97.2%; Score 10170; DB 1; Length 2621;

Best Local Similarity 96.7%; Pred. No. 0;  
 Matches 1919; Conservative 36; Mismatches 30; Indels 0; Gaps 0;  
 QY 1 MAPITAYSQOTRGLGCIITSLTGRDNOVEGEVQVWSTATOSFLATCVCNGCWTVYHGA 60  
 Db 637 LAPITAYSQOTRGLGCIITSLTGRDKNQVEGEVQVWSTATOSFLATCVCNGCWTVYHGA 696  
 QY 61 GSKTLAGPKGITOMYTNVDQLVGMQAPPGARSFTCTCGSSDLVLRHADVIPVRRR 120  
 Db 697 GSKTLAAPKGPITOMYTNVDQLVGMQAPPGARSFTCTCGSSDLVLRHADVIPVRRR 756  
 QY 121 GDSRGLSLSPRVSYLKGSSGGLLCPSCSHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 180  
 Db 757 GDSRGLSLSPRVSYLKGSSGGLLCPFGHAVGIFRAAVCTRGVAKAVDFVPVSEMETTM 816  
 QY 181 RSPVFTDNSSPPPAVPOTFQVAHLHAPTGSKSTKVAFAAQAQYKVLNLPNSVAATLFGF 240  
 Db 817 RSPVFTDNSSPPPAVPOSFQVAHLHAPTGSKSTKVAFAAQAQYKVLNLPNSVAATLFGF 876  
 QY 241 AYMSKAHGIDPNIRTVRTITTCAPITYSTYTGKFLADGGCGGAYDIIICDECHSTDSTT 300  
 Db 877 AYMSKAHGIDPNIRTVRTITTCAPITYSTYTGKFLADGGCGGAYDIIICDECHSTDSTT 936  
 QY 301 ILGIGTVLQDAETAGARLVVLAATATPPGVTVPHPNIEEVALSGTGEIPYGAIPETI 360  
 Db 937 ILGIGTVLQDAETAGARLVVLAATATPPGVTVPHPNIEEVALSGTGEIPYGAIPETI 996  
 QY 361 KGRHLIFCHSKKKCDLAAKSLGNGAVAYRGLDVSIVPTSGDVIVVATDALMTGFT 420  
 Db 997 RGRHLIFCHSKKKCDLAAKSLGNGAVAYRGLDVSIVPTSGDVIVVATDALMTGFT 1056  
 QY 421 GDFDSVIDCNTCTVTQVDFSLDPTFTIETTTVPQDAVRSORRGRTGRGMGIYRFVTPG 480  
 Db 1057 GDFDSVIDCNTCTVTQVDFSLDPTFTIETTTVPQDAVRSORRGRTGRGMGIYRFVTPG 1116  
 QY 481 ERPSGMPDSSVLCEDYDAGCAWYELTAPETSVELRAYLNTPLPVCODHLEFWSVETGL 540  
 Db 1117 ERPSGMPDSSVLCEDYDAGCAWYELTAPETSVELRAYLNTPLPVCODHLEFWSVETGL 1176  
 QY 541 THIDAHELSTKQAGDNFPYLVAQATVCARAQAPPSWQMKLRLKPTLHGPTPL 600  
 Db 1177 THIDAHELSTKQAGDNFPYLVAQATVCARAQAPPSWQMKLRLKPTLHGPTPL 1236  
 QY 601 YRLGAVQNEVTTTHPTIKYIMACMSADLEVTTSTWLVGVLAAALAYCITGTSWIVGR 660  
 Db 1237 YRLGAVQNEVTLTHPTIKYIMACMSADLEVTTSTWLVGVLAAALAYCITGTSWIVGR 1296  
 QY 661 ILSGKPAIIPDREVLVREFDEMBECASHLPYIEQGMQLAEQFKQKALLOTATKQAEA 720  
 Db 1297 ILSGKPAIIPDREVLVREFDEMBECASHLPYIEQGMQLAEQFKQKALLOTATKQAEA 1356  
 QY 721 AAPVVEKWRITLFAFWAKHWNFTISGLOYLGLSTLPGNPAIASIMAFASITSPITLTHQ 780  
 Db 1357 AAPVVEKWRITLFAFWAKHWNFTISGLOYLGLSTLPGNPAIASIMAFASITSPITLTHQ 1416  
 QY 781 TLLFNILGGWVAQAOLAPPASAFAVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
 Db 1417 TLLFNILGGWVAQAOLAPPASAFAVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 1476  
 QY 841 PKVMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAGVWNNRLIAFASR 900  
 Db 1477 PKVMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAGVWNNRLIAFASR 1536  
 QY 901 GNHVSPTHYVPESDAAARVTOILSSLTITQLKRLHOWINEDCTPCSGSWLRDWDWIC 960  
 Db 1537 GNHVSPTHYVPESDAAARVTOILSSLTITQLKRLHOWINEDCTPCSGSWLRDWDWIC 1596  
 QY 961 TVLTDFKTLQSKLLPLPGVPFFSCQSGYKVGWRGDMOTTCPCGAIQTHGVKNGSMR 1020  
 Db 1597 TVLTDFKTLQSKLLPLPGVPFFSCQSGYKVGWRGDMOTTCPCGAIQTHGVKNGSMR 1656  
 QY 1021 IVGPRTCSNTHGTFPINAATTGCTPSPAPNYSRALWRVAEEYVEVTRVGDHFYVTGM 1080

Db 1657 IVGPKTCSNTWHGTFPPINAYTTGCTPTSPAPNTSRALWRVAEEYVEVTRVGDHFHYTGM 1716  
QY 1081 TTDNVKPCQVPAPPEFTVDGVRILHRYAPACKPLREBEVTFVLGNQYLVGSQLCEPE 1140  
Db 1717 TTDNVKPCQVPAPPEFTVDGVRILHRYAPACKPLREBEVTFVLGNQYLVGSQLCEPE 1776  
QY 1141 PDVAVLTSMITDPSHITATKARRLARGSPPLSSASSASQLSAPSKATCTTHHSDPDAD 1200  
Db 1777 PDVAVLTSMITDPSHITATKARRLARGSPPLSSASSASQLSAPSKATCTTHHSDPDAD 1836  
QY 1201 LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSYPAEILRRSRKPPRAM 1260  
Db 1837 LIEANLLWQEMGNITRVESENKVVILDSFEPLQAEEDEREVSYPAEILRRSRKPPRAM 1896  
QY 1261 PIWARDYNPELLESKWDPPYVPPVHGCPLPPAKAPPIPPRRKRTTVLSSTVSSALA 1320  
Db 1897 PIWARDYNPELLESKWDPPYVPPVHGCPLPPAKAPPIPPRRKRTTVLSSTVSSALA 1956  
QY 1321 ELATKTFGSSSADVSGTATASPDQPSDDGAGSDVESYSSMPLLEGPGDPDLSDGSW 1380  
Db 1957 ELATKTFGSSSADVSGTATALPDQASDDGDKGSDVESYSSMPLLEGPGDPDLSDGSW 2016  
QY 1381 STVSEBESDVCCSMSTWTGALITPCAEEETKLPINALSNSLLRHNNLVVATTSRAS 1440  
Db 2017 STVSEBESDVCCSMSTWTGALITPCAEEETKLPINALSNSLLRHNNLVVATTSRAS 2076  
QY 1441 LRQKVTFRDQLVDDHVDLKENKAKASTVKALLSVEEACKLTPHSAKSKFGYAK 1500  
Db 2077 LRQKVTFRDQLVDDHVDLKENKAKASTVKALLSVEEACKLTPHSAKSKFGYAK 2136  
QY 1501 DVRLSSKAVNHISVWKDLLEDTPIDTITMAKNEVFCVQPEKGGKPKARLIVFPDLG 1560  
Db 2137 DVRLSSKAVNHISVWKDLLEDTPIDTITMAKNEVFCVQPEKGGKPKARLIVFPDLG 2196  
QY 1561 VRVCEKALYDVWSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMFAIDTRCFDS 1620  
Db 2197 VRVCEKALYDVWSTLPQAVMGSSYGFQYSPQORVEFLVNAWAKKCPMFAIDTRCFDS 2256  
QY 1621 TVTENDIRVESIYQCCDLAPEARQAIISLTERLYIGGPLTNSKQNGCYRRCRASGVL 1680  
Db 2257 TVTENDIRVESIYQCCDLAPEARQAIISLTERLYIGGPLTNSKQNGCYRRCRASGVL 2316  
QY 1681 TSCGNTLTCYLAACRAAKLQDCTMLVCGDDLVICESAGTORDEASLRAFTEAMTRY 1740  
Db 2317 TSCGNTLTCYLAACRAAKLQDCTMLVCGDDLVICESAGTORDEASLRAFTEAMTRY 2376  
QY 1741 SAPCPDPKPYDLELITSCSSNVSVADHAGKRVYLTTRDPTTPLAFAAWETARHTPVN 1800  
Db 2377 SAPCPDPKPYDLELITSCSSNVSVADHAGKRVYLTTRDPTTPLAFAAWETARHTPVN 2436  
QY 1801 SWLGNIIMYAPTLWARMILTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIIRL 1860  
Db 2437 SWLGNIIMYAPTLWARMILTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIIRL 2496  
QY 1861 HGLSAPFSLHSPGEINRVASCLKGLVPLVWRHARSVRARLLSQGGRAATCGKYL 1920  
Db 2497 HGLSAPFSLHSPGEINRVASCLKGLVPLVWRHARSVRARLLSQGGRAATCGKYL 2556  
QY 1921 NWAIVTKLTPPAAQSLDSSWFAVYGGDIYHLSLRARPFWFMCILLLSVGVGIY 1980  
Db 2557 NWAIVTKLTPPAAQSLDSSWFAVYGGDIYHLSLRARPFWFMCILLLSVGVGIY 2616  
QY 1981 LLPNR 1985  
Db 2617 LLPNR 2621

RESULT 15  
US-08-384-616-36  
; Sequence 36, Application US/08384616  
; Patent No. 5847101  
; GENERAL INFORMATION:  
; APPLICANT: OKAYAMA, Hiroto

APPLICANT: FUKU, Isao  
APPLICANT: MORI, Chisato  
APPLICANT: TAKAMIZAWA, Akahisa  
APPLICANT: YOSHIDA, Iwao  
TITLE OF INVENTION: NON-A, NON-B HEPATITIS VIRUS GENOMIC  
TITLE OF INVENTION: CDNA AND ANTIGEN POLYPEPTIDE  
NUMBER OF SEQUENCES: 50  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Armstrong, Westerman, Hattori, McLeeland &  
ADDRESSEE: Naughton  
STREET: 1725 K St. N.W. Suite 1000  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
ZIP: 20006  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.5 in, 1.44Mb  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS, Version 5.0  
SOFTWARE: ASCII  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/384,616  
FILING DATE:  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/769,996  
FILING DATE: 02-OCT-1991  
APPLICATION NUMBER: JP 2-167466  
FILING DATE: 25-JUN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-230921  
FILING DATE: 31-AUG-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 2-305605  
FILING DATE: 09-NOV-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/635,451  
FILING DATE: 28-DEC-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens-Smith, Theresa M.  
REGISTRATION NUMBER: 36,281  
REFERENCE/DOCKET NUMBER: 900703B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 659-2930  
TELEFAX: (202) 887-0357  
TELEX: 440142  
INFORMATION FOR SEQ ID NO: 36:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2621 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-384-616-36

Query Match 97.2%; Score 10170; DB 2; Length 2621;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 1919; Conservative 36; Mismatches 30; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLGCIITSLTGRDNQVEGEVQVYSTATQSFATCVNGVCMVYHGA 60  
Db 637 LAPITAYSQOTRGLGCIITSLTGRDNQVEGEVQVYSTATQSFATCVNGVCMVYHGA 696  
QY 61 GSKTLGAPGPIQMTYNVDQDLVGHQAPGARSLLTCTGSSDYLVTTHADVIPVRRR 120  
Db 697 GSKTLGAPGPIQMTYNVDQDLVGHQAPGARSLLTCTGSSDYLVTTHADVIPVRRR 756  
QY 121 GDSRGLSPRPVSYLKGSGGFLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 757 GDSRGLSPRPVSYLKGSGGFLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 816  
QY 181 RSPVFTDNSPPAVPOTFOVAHLHAPTGSCKSTKVPAAVAAQGYKVLINPNSVAATLGF 240  
Db 817 RSPVFTDNSPPAVPOTFOVAHLHAPTGSCKSTKVPAAVAAQGYKVLINPNSVAATLGF 876

QY	241	AYMSKAHGDIPNIRTVRTTTCAPITYSTYKFIADGCGSGGAYDIIICDECHSTDSTT	300
Db	877	AYMSKAHGDIPNIRTVRTTTCAPITYSTYKFIADGCGSGGAYDIIICDECHSTDSTT	936
QY	301	ILGIGTVLDQAEATAGARLVVLAATATPPGVTVPHPNIEVALSSTGEIPIFYKAIPIETI	360
Db	937	ILGIGTVLDQAEATAGARLVVLAATATPPGVTVPHPNIEVALSSTGEIPIFYKAIPIEAI	996
QY	361	KGRGHLIFCHSKKKCDLAALKSLGGLNAVAYYRGDLVSVIPISGDIIVVATDALTMTGPT	420
Db	997	KGRGHLIFCHSKKKCDLAALKSLGGLNAVAYYRGDLVSVIPITGDVWVATDALTMTGYT	1056
QY	421	GFDSVIDONTCTVQTVDFSLDPTFTIETTVPQDAVSRSQRGRGGRGMYRVFTPG	480
Db	1057	GFDSVIDONTCTVQTVDFSLDPTFTIETTVPQDAVSRSQRGRGGRGMYRVFTPG	1116
QY	481	ERBSGMFDSVLCCEYDAGCAWVELTPAETSVELRAYLNTPLPVCQDHLFEFVESFTGL	540
Db	1117	ERBSGMFDSVLCCEYDAGCAWVELTPAETSVELRAYLNTPLPVCQDHLFEFVESFTGL	1176
QY	541	THIDAHFLSOTKQAGDNFPYIVAYQATVCARAQAPPSWDQMKCLIRLKTTLHGPTPLL	600
Db	1177	THIDAHFLSOTKQAGDNFPYIVAYQATVCARAQAPPSWDQMKCLIRLKTTLHGPTPLL	1236
QY	601	YRLGAVONEVTHPTITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSMVIVGR	660
Db	1237	YRLGAVONEVTHPTITKYIMACMSADLEVTSTWLVGGVLAALAAAYCLTTGSMVIVGR	1296
QY	661	IILSGKALITPDRVELVREDEMEECASHLPYIEQGMQLAEQFKQKALIGLLOATKQAEA	720
Db	1297	IILSGKALITPDRVELVREDEMEECASHLPYIEQGMQLAEQFKQKALIGLLOATKQAEA	1356
QY	721	AAPVVEKWRITLFAFWAKHWNFTSGIYLAGLSTLFCNPAIASLMAFTASITSLPTQH	780
Db	1357	AAPVVEKWRITLFAFWAKHWNFTSGIYLAGLSTLFCNPAIASLMAFTASITSLPTQH	1416
QY	781	TLLNIIIGWVAQAAPSAASAFVAGIAGAAVGSIGLKVLDIILAGYAGAGALVA	840
Db	1417	TLLNIIIGWVAQAAPSAASAFVAGIAGAAVGSIGLKVLDIILAGYAGAGALVA	1476
QY	841	FKWMSGMPSTEDLVNLLPALSPGALVGVVCAAILRRHVGPGEQAVQWNNRLIAFASR	900
Db	1477	FKWMSGMPSTEDLVNLLPALSPGALVGVVCAAILRRHVGPGEQAVQWNNRLIAFASR	1536
QY	901	GNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHWINEDCSTPCSGMWLRDWDWIC	960
Db	1537	GNHVSPTHYVPESDAAARVTQILSSLTITQLKRLHWINEDCSTPCSGMWLRDWDWIC	1596
QY	961	TVLTDFTKWLQSKLLPRLPGVPPFSCQGYKGVWRGDIMQTTCPGQAQITGHVKNQSMR	1020
Db	1597	TVLTDFTKWLQSKLLPRLPGVPPFSCQGYKGVWRGDIMQTTCPGQAQITGHVKNQSMR	1656
QY	1021	IVGPRTCNTWHGTFFPINAYTTGCTPSPAPNYGRALMRVAEEVEVTRVGDHYVTGM	1080
Db	1657	IVGPRTCNTWHGTFFPINAYTTGCTPSPAPNYGRALMRVAEEVEVTRVGDHYVTGM	1716
QY	1081	TTDNVKKCPQVPAPEFFTEVDGVRILHRVAPACKPLLRREEVFLVGLNQYLVGSLPCBPE	1140
Db	1717	TTDNVKKCPQVPAPEFFTEVDGVRILHRVAPACKPLLRREEVFLVGLNQYLVGSLPCBPE	1776
QY	1141	PDVAVLTSMLTDPSSHITAEATARRLARGSPSLASSASQLSAPLSKATCTTRHDSPPAD	1200
Db	1777	PDVAVLTSMLTDPSSHITAEATARRLARGSPSLASSASQLSAPLSKATCTTRHDSPPAD	1836
QY	1201	LIEANLLWRQEMGNIITRVSEKNVILDSFPELQAEEDEREVSVPAEILRKRKFPFRAM	1260
Db	1837	LIEANLLWRQEMGNIITRVSEKNVILDSFPELQAEEDEREVSVPAEILRKRKFPFRAM	1896
QY	1261	PIWAPDPYNPPLLESWKOPDYVPPVHVHGCPLPPAKAPPIPPPRRKRRTVVLSESTVSSALA	1320
Db	1897	PIWAPDPYNPPLLESWKOPDYVPPVHVHGCPLPPAKAPPIPPPRRKRRTVVLSESTVSSALA	1956

Search completed: December 8, 2004, 12:33:52  
Job time : 43 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:25:07 ; Search time 129 Seconds

(without alignments)  
5496.116 Million cell updates/sec

Title: US-09-576-989-3

Perfect score: 10465

Sequence: 1 MAPITAYSQTRGLGCIIT.....FMWCLLLSVGVIGVILLFNR 1985

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1585576 seqs, 357178320 residues

Total number of hits satisfying chosen parameters: 1585576

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

#### Database :

Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US10E\_PUBCOMB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Query Score	Match	Length	DB ID	Description
1	10465	100.0	1985	16	US-10-639-150-2
2	10462	100.0	3010	15	US-10-467-000-1
3	10459	99.9	1985	14	US-10-259-275-42
4	10442	99.8	2201	13	US-10-029-907-3
5	10442	99.8	2201	14	US-10-309-561-3
6	10442	99.8	2201	16	US-10-789-355-3
7	10442	99.8	2201	17	US-10-686-835-3
8	10170	97.2	2201	13	US-10-085-476-2
9	10157	97.1	3010	16	US-10-333-449A-34
10	9710	92.8	2985	14	US-10-259-275-40
11	9454.5	90.3	3011	15	US-10-296-734-406
12	9434.5	90.2	3011	9	US-09-742-659-4
13	9430.5	90.1	3011	10	US-09-891-894-3
					Sequence 2, Appli
					Sequence 1, Appli
					Sequence 42, Appl
					Sequence 3, Appli
					Sequence 3, Appli
					Sequence 3, Appli
					Sequence 2, Appli
					Sequence 34, Appl
					Sequence 40, Appl
					Sequence 406, App
					Sequence 4, Appli
					Sequence 3, Appli

14	9430.5	90.1	3011	14	US-10-184-150-3	Sequence 3, Appli
15	9430.5	90.1	3011	14	US-10-328-997-3	Sequence 3, Appli
16	9430.5	90.1	3012	9	US-09-238-076-2	Sequence 2, Appli
17	9430.5	90.1	3012	10	US-09-995-937-2	Sequence 2, Appli
18	9430.5	90.1	3012	10	US-09-917-583-2	Sequence 2, Appli
19	9418.5	90.0	3011	9	US-09-952-572-9	Sequence 9, Appli
20	9418.5	90.0	3011	9	US-09-747-419-20	Sequence 20, Appl
21	9418.5	90.0	3011	14	US-10-259-275-20	Sequence 20, Appl
22	9418.5	90.0	3011	15	US-10-189-359-14	Sequence 14, Appl
23	9417.5	90.0	3011	9	US-09-916-359-2	Sequence 2, Appli
24	9417.5	90.0	3011	16	US-10-445-724-2	Sequence 2, Appli
25	9371.5	89.6	3011	9	US-09-238-076-20	Sequence 20, Appl
26	9371.5	89.6	3011	10	US-09-995-937-20	Sequence 20, Appl
27	9371.5	89.6	3011	10	US-09-917-563-20	Sequence 20, Appl
28	9324.5	89.1	3011	14	US-10-232-643-6	Sequence 6, Appli
29	9266.5	88.5	3011	9	US-09-929-955-1	Sequence 1, Appli
30	9266.5	88.5	3011	13	US-10-104-966-1	Sequence 1, Appli
31	9266.5	88.5	3011	15	US-10-719-619-1	Sequence 1, Appli
32	9266.5	88.5	3011	17	US-10-817-591-1	Sequence 1, Appli
33	8799.5	84.1	2894	9	US-09-941-611-23	Sequence 23, Appl
34	8799.5	84.1	2894	14	US-10-044-995-23	Sequence 23, Appl
35	7638.5	73.0	2940	14	US-10-226-629A-13	Sequence 13, Appl
36	7168	68.5	1736	14	US-10-328-127-2	Sequence 2, Appli
37	7168	68.5	1736	14	US-10-328-206-2	Sequence 2, Appli
38	4823.5	46.1	2865	9	US-09-742-659-6	Sequence 6, Appli
39	4725.5	45.2	1040	9	US-09-929-955-9	Sequence 9, Appli
40	4725.5	45.2	1040	13	US-10-104-966-9	Sequence 9, Appli
41	4725.5	45.2	1040	15	US-10-719-619-9	Sequence 9, Appli
42	4725.5	45.2	1040	17	US-10-817-591-9	Sequence 9, Appli
43	4443	42.5	2307	10	US-09-919-301-2	Sequence 2, Appli
44	4443	42.5	2307	14	US-10-191-966-2	Sequence 2, Appli
45	4440	42.4	2307	10	US-09-919-301-9	Sequence 9, Appli

#### ALIGNMENTS

#### RESULT 1

US-10-639-150-2  
; Sequence 2, Application US/10639150  
; Publication No. US20040121975A1  
; GENERAL INFORMATION:  
; APPLICANT: BRISTOL-MYERS SQUIBB COMPANY  
; TITLE OF INVENTION: HEPATITIS C VIRUS ASSAYS  
; FILE REFERENCE: D0224 NP  
; CURRENT APPLICATION NUMBER: US/10/639,150  
; CURRENT FILING DATE: 2003-08-12  
; PRIOR APPLICATION NUMBER: US 60/402,661  
; PRIOR FILING DATE: 2002-08-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 2  
; LENGTH: 1985  
; TYPE: PRT  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: HCV Replicon  
US-10-639-150-2

Query Match	100.0%;	Score 10465;	DB 16;	Length 1985;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1985;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MAPITAYSQTRGLGCIITSLTGRDRNOVEGEVQVVSATQSFATCVNGVCTVYVHGA	60	
Db	1	MAPITAYSQTRGLGCIITSLTGRDRNOVEGEVQVVSATQSFATCVNGVCTVYVHGA	60	
Qy	61	GSKTLGAPKGFITQMTYNVDQDLVGWQAPPGARSLLPCTCGSSDLVLTTRHADVIPVRRR	120	
Db	61	GSKTLGAPKGFITQMTYNVDQDLVGWQAPPGARSLLPCTCGSSDLVLTTRHADVIPVRRR	120	
Qy	121	GDSSGSLSPRPVSYLKGSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM	180	

Db 121 GDSRGLSPRVSVYLKSGSGPLLCPSGHAVGI FRAAVCTRGVAKAVDFVPVESMETTM 180  
 QY 181 RSPVFTDNSSPPAVPQTQVAVHLHAPTSGKSTKVPAAAYAAQYKVLVNLNPSVAATLGF 240  
 Db 181 RSPVFTDNSSPPAVPQTQVAVHLHAPTSGKSTKVPAAAYAAQYKVLVNLNPSVAATLGF 240  
 QY 241 AYSKAGHDNIRTOVRTITTCAPITYTYGKFLADGCGSGAYDIIICDECHSDSTT 300  
 Db 241 AYSKAGHDNIRTOVRTITTCAPITYTYGKFLADGCGSGAYDIIICDECHSDSTT 300  
 QY 301 ILGIGTVLDOAETAGARLVVLAATAPGSGVTVPHPNIEEVALSGTGEIPYGAIPETI 360  
 Db 301 ILGIGTVLDOAETAGARLVVLAATAPGSGVTVPHPNIEEVALSGTGEIPYGAIPETI 360  
 QY 361 KGRHLIFCHSKKKKDELAALKSGLGLNAVAYYRGDLVSVIPTSGDVIVVATDALMTGFT 420  
 Db 361 KGRHLIFCHSKKKKDELAALKSGLGLNAVAYYRGDLVSVIPTSGDVIVVATDALMTGFT 420  
 QY 421 GDFDSVIDCNTCTVOTVDFSLDPTFTTETTTVPQDAVSRSORGRGGRMGYIRFVTPG 480  
 Db 421 GDFDSVIDCNTCTVOTVDFSLDPTFTTETTTVPQDAVSRSORGRGGRMGYIRFVTPG 480  
 QY 481 ERSFGMFDSSVCECYDAGCAWYELTPAETSRLRAYLNTPLPVCQDHLFEFVESFTGL 540  
 Db 481 ERSFGMFDSSVCECYDAGCAWYELTPAETSRLRAYLNTPLPVCQDHLFEFVESFTGL 540  
 QY 541 THIDAHLFSTQKAGDNFPYLVAYQAVTCARAQAPPPSWDMKCLIRLKPTELHGTPL 600  
 Db 541 THIDAHLFSTQKAGDNFPYLVAYQAVTCARAQAPPPSWDMKCLIRLKPTELHGTPL 600  
 QY 601 YRLGAVONEVTTTHPIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSGVTVGR 660  
 Db 601 YRLGAVONEVTTTHPIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSGVTVGR 660  
 QY 661 IILSGKPAIIPDREVLRYEFDEMEBCASHLPYIEQGMQLAEQFKQKAIQLLOTATKQAEA 720  
 Db 661 IILSGKPAIIPDREVLRYEFDEMEBCASHLPYIEQGMQLAEQFKQKAIQLLOTATKQAEA 720  
 QY 721 AAPVVEKWTLEAFWAKHWNFTSGTQYLAGLSTLPGNPAIASLMAFTASITSPLTQ 780  
 Db 721 AAPVVEKWTLEAFWAKHWNFTSGTQYLAGLSTLPGNPAIASLMAFTASITSPLTQ 780  
 QY 781 TLLFNILGGWVAQALAPPSAASAVGAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
 Db 781 TLLFNILGGWVAQALAPPSAASAVGAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
 QY 841 FKWMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVOMNRLIAFASR 900  
 Db 841 FKWMSGEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVOMNRLIAFASR 900  
 QY 901 GNVSPHYVPESDAAARVTOILSSLITQLLKLHOWINEDCSTPCSGSWLRDWDWIC 960  
 Db 901 GNVSPHYVPESDAAARVTOILSSLITQLLKLHOWINEDCSTPCSGSWLRDWDWIC 960  
 QY 961 TVLDEKTLQSKLLPRLPGVPFFSCORGKGVWRGDMOTTCPCQAQITGHVKGSMR 1020  
 Db 961 TVLDFKTLQSKLLPRLPGVPFFSCORGKGVWRGDMOTTCPCQAQITGHVKGSMR 1020  
 QY 1021 IVGPRTCSNTHWGTFFPINAYTTGCTPSPAPNYGRALWRVAEEVVEVTRVGDHYVTGM 1080  
 Db 1021 IVGPRTCSNTHWGTFFPINAYTTGCTPSPAPNYGRALWRVAEEVVEVTRVGDHYVTGM 1080  
 QY 1081 TTDNVKPCQVPAPEFFTEVDGVRILHRYPACKLLREVTFLVGLNQLVGSOLPCEPE 1140  
 Db 1081 TTDNVKPCQVPAPEFFTEVDGVRILHRYPACKLLREVTFLVGLNQLVGSOLPCEPE 1140  
 QY 1141 PDVAVLTSMLTDPGSHITAEAKRLARGSPSLASSASQLSAPSLKATCTTRHSDPAD 1200  
 Db 1141 PDVAVLTSMLTDPGSHITAEAKRLARGSPSLASSASQLSAPSLKATCTTRHSDPAD 1200  
 QY 1201 LIEANLLWRQBMGNITRVSEKNVILDSPEPLQAEDEDEVSPABILRRSKRFFPRAM 1260  
 Db 1201 LIEANLLWRQBMGNITRVSEKNVILDSPEPLQAEDEDEVSPABILRRSKRFFPRAM 1260

QY 1261 PIWARPDPNPPLESKWDPDYVPPVVGCPPLPAKAPPIPPRRKKTIVLSESTVSALA 1320  
 Db 1261 PIWARPDPNPPLESKWDPDYVPPVVGCPPLPAKAPPIPPRRKKTIVLSESTVSALA 1320  
 QY 1321 ELATKTFGSSSAVDSGTATASPDQSDDGAGSDVESYSSMPPLEGPGDPLSDGWS 1380  
 Db 1321 ELATKTFGSSSAVDSGTATASPDQSDDGAGSDVESYSSMPPLEGPGDPLSDGWS 1380  
 QY 1381 STVSEBASDVVCCSMSTYTWTCALITPCAABETKLPINALSNLLRHHNLVYATTSRSAS 1440  
 Db 1381 STVSEBASDVVCCSMSTYTWTCALITPCAABETKLPINALSNLLRHHNLVYATTSRSAS 1440  
 QY 1441 LRQKVTYFRLQVLDHVDVLEKEMKAKASTVAKALLSVEEACKLTPPHSARSKFYGA 1500  
 Db 1441 LRQKVTYFRLQVLDHVDVLEKEMKAKASTVAKALLSVEEACKLTPPHSARSKFYGA 1500  
 QY 1501 DVNLSKAVNHIRSVMKOLLEDTETPTTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
 Db 1501 DVNLSKAVNHIRSVMKOLLEDTETPTTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
 QY 1561 VRVCEKALYDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
 Db 1561 VRVCEKALYDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
 QY 1621 TVTENDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGLTNSKGONCGYRRCRASGVL 1680  
 Db 1621 TVTENDIRVEESIYQCCDLAPPEARQAIRSLTERLYIGGLTNSKGONCGYRRCRASGVL 1680  
 QY 1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEDEASLRAFTTEAMTRY 1740  
 Db 1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEDEASLRAFTTEAMTRY 1740  
 QY 1741 SAPGDPPEPEYDLELITSCSSNVSVAHADSGKRVYVLTDRPTPLARAWEATARHTPVN 1800  
 Db 1741 SAPGDPPEPEYDLELITSCSSNVSVAHADSGKRVYVLTDRPTPLARAWEATARHTPVN 1800  
 QY 1801 SWLGNIMYAPTILWARMILMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIIRL 1860  
 Db 1801 SWLGNIMYAPTILWARMILMTHFFSILLAOLEKALDCQIYGACYSIEPLDLPQIIRL 1860  
 QY 1861 HGLSFLSHSYSGEINRVASCLRLKGLVPLRVRHRRARSVRARLLSQGGRATCGKYL 1920  
 Db 1861 HGLSFLSHSYSGEINRVASCLRLKGLVPLRVRHRRARSVRARLLSQGGRATCGKYL 1920  
 QY 1921 NWAVRTKLTLPAPASQDLSSWVAGYSGDIIYHLSRARPRWFWMCILLISVGVGIY 1980  
 Db 1921 NWAVRTKLTLPAPASQDLSSWVAGYSGDIIYHLSRARPRWFWMCILLISVGVGIY 1980  
 QY 1981 LLPNR 1985  
 Db 1981 LLPNR 1985

RESULT 2  
 US-10-467-000-1  
 ; Sequence 1, Application US/10467000  
 ; Publication No. US20040067486A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: De Francesco, Raffaele  
 ; APPLICANT: Migliaccio, Giovanni  
 ; APPLICANT: Paonessa, Giacomo  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS REPLICONS AND REPLICON  
 ; TITLE OF INVENTION: ENHANCED CELLS  
 ; FILE REFERENCE: ITR0003P  
 ; CURRENT APPLICATION NUMBER: US/10/467,000  
 ; CURRENT FILING DATE: 2003-07-21  
 ; PRIOR APPLICATION NUMBER: PCT/EP02/00526  
 ; PRIOR FILING DATE: 2002-01-16  
 ; PRIOR APPLICATION NUMBER: 60/263,479  
 ; PRIOR FILING DATE: 2001-01-23  
 ; NUMBER OF SEQ ID NOS: 13  
 ; SOFTWARE: FastSEQ for Windows Version 4.0



```
; SEQ ID NO 1
; LENGTH: 3010
; TYPE: PRT
; ORGANISM: Con 1 HCV isolate nucleic acid
US-10-467-000-1

Query Match      100.0%; Score 10462; DB 15; Length 3010;
Best Local Similarity 99.9%; Pred No. 0;
Matches 1984; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQFLATCVNGVCMVYHGA 60
Db :
1026 LAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQFLATCVNGVCMVYHGA 1085

QY 61 GSKTLAGPKGPIQMYTNVDQDLVGMQAPPAGARSITPCTCGSSDLYLVTRHADVIPVRRR 120
Db :
1086 GSKTLAGPKGPIQMYTNVDQDLVGMQAPPAGARSITPCTCGSSDLYLVTRHADVIPVRRR 1145

QY 121 GDSRGSLLSPRPVSYLKSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180
Db :
1146 GDSRGSLLSPRPVSYLKSSGGPILCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205

QY 181 RSPVFTDNSSPPAVPQTQVAHLHAPTSGKSTKVPAAYAAQGYKVLVNFPSVAATLFGF 240
Db :
1206 RSPVFTDNSSPPAVPQTQVAHLHAPTSGKSTKVPAAYAAQGYKVLVNFPSVAATLFGF 1265

QY 241 AYMSKAHGIDENIRGVRTITTGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 300
Db :
1266 AYMSKAHGIDENIRGVRTITTGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 1325

QY 301 ILGIGTVLDAQETAGARLVWLATATPPGVSIVPHNIEEVALSSTGEIPFGKAIPIETI 360
Db :
1326 ILGIGTVLDAQETAGARLVWLATATPPGVSIVPHNIEEVALSSTGEIPFGKAIPIETI 1385

QY 361 KGGHILIPCHSKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIWATDALMTGFT 420
Db :
1386 KGGHILIPCHSKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIWATDALMTGFT 1445

QY 421 GDFSVDICNTCVTQVDFSLDPTFTTITTTVPQDAVRSORRGRTGGRGVIYRFTVPG 480
Db :
1446 GDFSVDICNTCVTQVDFSLDPTFTTITTTVPQDAVRSORRGRTGGRGVIYRFTVPG 1505

QY 481 ERPSGMFDSSVLCBCEYDAGCAWYELTPAETSVRLAYLINTPLPVCODHLEFWSVFGL 540
Db :
1506 ERPSGMFDSSVLCBCEYDAGCAWYELTPAETSVRLAYLINTPLPVCODHLEFWSVFGL 1565

QY 541 THIDAHFLSOTKQAGDNFPYLVAQVATVCARAQAPPPSDQMWKCLIRLKPILHGPITPL 600
Db :
1566 THIDAHFLSOTKQAGDNFPYLVAQVATVCARAQAPPPSDQMWKCLIRLKPILHGPITPL 1625

QY 601 YRLGAVONEVTTTHPIITKYINACMSADLEVTTSTWLVGGVLAALAAAYCLTGSVVIAGR 660
Db :
1626 YRLGAVONEVTTTHPIITKYINACMSADLEVTTSTWLVGGVLAALAAAYCLTGSVVIAGR 1685

QY 661 IILSGKPAIIPDREVLVREFDEMEBCASHLPYIEQGMQLABQFQKQKALGLLOTATKQAEA 720
Db :
1686 IILSGKPAIIPDREVLVREFDEMEBCASHLPYIEQGMQLABQFQKQKALGLLOTATKQAEA 1745

QY 721 AAPVVEKWRITLEAFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTOH 780
Db :
1746 AAPVVEKWRITLEAFWAKHMMNFISGIQYLAGLSTLPGNPAIASLMAFTASITSPLTTOH 1805

QY 781 TLLFNILGGWVAAGLAPPSAASAFVAGIAGAAGVSGTGLKVLVDIILAGYGAGVAGALVA 840
Db :
1806 TLLFNILGGWVAAGLAPPSAASAFVAGIAGAAGVSGTGLKVLVDIILAGYGAGVAGALVA 1865

QY 841 FKVMGEMSPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQWNNRIIAFASR 900
Db :
1866 FKVMGEMSPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVQWNNRIIAFASR 1925

QY 901 GNHVSPTHYPESDAAARVTOILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC 960
Db :
1926 GNHVSPTHYPESDAAARVTOILSSLTITQLKRLHQWINEDECSTPCSGSWLRDWDWIC 1985
```

RESULT 3  
 US-10-259-275-42  
 ; Sequence 42, Application US/10259275  
 ; Publication No. US20030125541A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lemon, Stanley M.  
 ; APPLICANT: Y1, Minkyung  
 ; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE  
 ; FILE REFERENCE: 265.0007 0120  
 ; CURRENT APPLICATION NUMBER: US/10/259,275  
 ; CURRENT FILING DATE: 2003-01-13  
 ; PRIOR APPLICATION NUMBER: US 60/171,909  
 ; PRIOR FILING DATE: 1999-12-23  
 ; PRIOR APPLICATION NUMBER: US 09/747,419  
 ; PRIOR FILING DATE: 2000-12-23  
 ; PRIOR APPLICATION NUMBER: US 60/325,236  
 ; PRIOR FILING DATE: 2001-09-27  
 ; PRIOR APPLICATION NUMBER: US 60/338,123  
 ; PRIOR FILING DATE: 2001-11-13  
 ; NUMBER OF SEQ ID NOS: 73  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 42  
 ; LENGTH: 1985  
 ; TYPE: PRN  
 ; ORGANISM: ARTIFICIAL  
 ; FEATURE:  
 ; OTHER INFORMATION: amino acid sequence encoded by the nucleotides 2119-8073 of  
 ; OTHER INFORMATION: SEQ ID NO:41  
 US-10-259-275-42

Query Match 99.9%; Score 10459; DB 14; Length 1985;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 1984; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 MAPITAYSQOQRLGGLIISLTGRDRNOVEGEVQVQVSTQSFATCVNGVCWTVVHGA 60  
 DB 1 MAPITAYSQOQRLGGLIISLTGRDRNOVEGEVQVQVSTQSFATCVNGVCWTVVHGA 60  
 QY 61 GSKTLAGKPGFITQMYTNVDQDLVGVQAPPGARSLETCGSSDLVLTVRHADVPVRR 120  
 DB 61 GSKTLAGKPGFITQMYTNVDQDLVGVQAPPGARSLETCGSSDLVLTVRHADVPVRR 120  
 QY 121 GDSRGLSPVSYLKGSSGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPEVSMETM 180  
 DB 121 GDSRGLSPVSYLKGSSGGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPEVSMETM 180  
 QY 181 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLFG 240  
 DB 181 RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSGKSTKVPAAAYAAQGYKVLVLPNSVAATLFG 240  
 QY 241 AYMSKAHGIDPNIRGTVRTITGAPITVSTYKFLADGGCGGGAYDIIICDECHSTDT 300  
 DB 241 AYMSKAHGIDPNIRGTVRTITGAPITVSTYKFLADGGCGGGAYDIIICDECHSTDT 300  
 QY 301 ILGIGTVDQETAGARLWLATATPGSVTVPHNIEEVALSGTEIPFGKAIPIETI 360  
 DB 301 ILGIGTVDQETAGARLWLATATPGSVTVPHNIEEVALSGTEIPFGKAIPIETI 360  
 QY 361 KGGRHLLPCHSKKKKDELAALKSLGLNAVAYRGLDVSIVPTSGDVIWVATDALMTGFT 420  
 DB 361 KGGRHLLPCHSKKKKDELAALKSLGLNAVAYRGLDVSIVPTSGDVIWVATDALMTGFT 420  
 QY 421 GDFDSVIDCNTCVITQVDFSLDPTFTIETTVPQDAVSRSORRGTRGRMGIRFVFTPG 480  
 DB 421 GDFDSVIDCNTCVITQVDFSLDPTFTIETTVPQDAVSRSORRGTRGRMGIRFVFTPG 480  
 QY 481 ERPSGMFSSVLCBYDAGCAWYELTPAETSURLRAYLNTPLPVCQDHLFEWESVFTGL 540  
 DB 481 ERPSGMFSSVLCBYDAGCAWYELTPAETSURLRAYLNTPLPVCQDHLFEWESVFTGL 540  
 QY 541 THIDAHFLSQTQKAGDNPPYLVAQATVCARAQAPPPSWDMWKCLIRLKPRTLHGPTPL 600

DB 541 THIDAHFLSQTQKAGDNPPYLVAQATVCARAQAPPPSWDMWKCLIRLKPRTLHGPTPL 600  
 QY 601 YRLGAVQNEVTTHTPITTKYIMACMSADLEWVSTWLVGVLAALAAAYCLTTCGVSIVVGR 660  
 DB 601 YRLGAVQNEVTTHTPITTKYIMACMSADLEWVSTWLVGVLAALAAAYCLTTCGVSIVVGR 660  
 QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLYPIEQMQLAEQFKQKAIGLLOTATQAEA 720  
 DB 661 IILSGKPAIIPDREVLYREFDEMEECASHLYPIEQMQLAEQFKQKAIGLLOTATQAEA 720  
 QY 721 AAPVVESKWTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSPITQH 780  
 DB 721 AAPVVESKWTLEAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSPITQH 780  
 QY 781 TLLFNILGGWVAQAAPPSSAASAFVAGIAGAAGVSGILGKVLVDIILAGVAGVAGALVA 840  
 DB 781 TLLFNILGGWVAQAAPPSSAASAFVAGIAGAAGVSGILGKVLVDIILAGVAGVAGALVA 840  
 QY 841 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVOMNRLIAFASR 900  
 DB 841 FKVMGEMPESTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAVOMNRLIAFASR 900  
 QY 901 GNVSPHYVPESDAAARVTQILSSLTITQLLRLHWNINEDCSTPCSGSWLRDWDWIC 960  
 DB 901 GNVSPHYVPESDAAARVTQILSSLTITQLLRLHWNINEDCSTPCSGSWLRDWDWIC 960  
 QY 961 TVLTDFKTLQSKLLPRLPGVPFFCQGVKGVWRGDMGIMQTTCCGCAQITGHVNGSMR 1020  
 DB 961 TVLTDFKTLQSKLLPRLPGVPFFCQGVKGVWRGDMGIMQTTCCGCAQITGHVNGSMR 1020  
 QY 1021 IVGPRTCSNTHGCTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTVGVGFHYVTGM 1080  
 DB 1021 IVGPRTCSNTHGCTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTVGVGFHYVTGM 1080  
 QY 1081 TTDNVKPCQVPAPEFTEVDGVLRLHRYAPACKPLLEEVTVFLVGLNQYLVGSLPCEPE 1140  
 DB 1081 TTDNVKPCQVPAPEFTEVDGVLRLHRYAPACKPLLEEVTVFLVGLNQYLVGSLPCEPE 1140  
 QY 1141 PDVAVLTSMLTDSHITAEAKERLARGSPSSASSASOLAPSILKATCTTRHSDSPAD 1200  
 DB 1141 PDVAVLTSMLTDSHITAEAKERLARGSPSSASSASOLAPSILKATCTTRHSDSPAD 1200  
 QY 1201 LIEANILWROEMGNITRVESENKVVILDSFEPLQAEDEEREVSVPAAEILRRSKFFPRAM 1260  
 DB 1201 LIEANILWROEMGNITRVESENKVVILDSFEPLQAEDEEREVSVPAAEILRRSKFFPRAM 1260  
 QY 1261 PIWARPYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVILSESTVSALA 1320  
 DB 1261 PIWARPYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVILSESTVSALA 1320  
 QY 1321 ELATKTFGSSSSAVDSGTATASDPQSDGSDGSDVESYSSMPLEGEPCDPLSDGWS 1380  
 DB 1321 ELATKTFGSSSSAVDSGTATASDPQSDGSDGSDVESYSSMPLEGEPCDPLSDGWS 1380  
 QY 1381 STVSEASESDVCCSMSTWTGALITPCAAEETKLIPINALSNSLLRHNILVYATTSRAS 1440  
 DB 1381 STVSEASESDVCCSMSTWTGALITPCAAEETKLIPINALSNSLLRHNILVYATTSRAS 1440  
 QY 1441 LRQKVVTFDLRLQVLDHRYOVLKEMKAKASTVKAKLLSVEEACKLTPHARSFEGYAK 1500  
 DB 1441 LRQKVVTFDLRLQVLDHRYOVLKEMKAKASTVKAKLLSVEEACKLTPHARSFEGYAK 1500  
 QY 1501 DVRLNSSKAVNHHSVWKDLLEDTEPIDTIMAKNEVFCVQPEKGRKPARLLVFPDLG 1560  
 DB 1501 DVRLNSSKAVNHHSVWKDLLEDTEPIDTIMAKNEVFCVQPEKGRKPARLLVFPDLG 1560  
 QY 1561 VRVCEKVALXDVWSTLFPQAVMGSSYGFQYSGQVFEFLVNAWAKKCPMGFAYDTRCFDS 1620  
 DB 1561 VRVCEKVALXDVWSTLFPQAVMGSSYGFQYSGQVFEFLVNAWAKKCPMGFAYDTRCFDS 1620  
 QY 1621 TVTENDIRVEESIYQCCDLAPAEARQATSLTERLYIGGPLTNSKONCGYRRCRASGVLT 1680  
 DB 1621 TVTENDIRVEESIYQCCDLAPAEARQATSLTERLYIGGPLTNSKONCGYRRCRASGVLT 1680

QY	1681	TSCGNTLTCYLKAAACRAAKLODCTMLVCGDGLVVICESACTOEBEASIRAFTEAMTRY	1740
Db	1681	TSCGNTLTCYLKAAACRAAKLODCTMLVCGDGLVVICESACTOEBEASIRAFTEAMTRY	1740
QY	1741	SAPPDPPKPEYDLELITSCSNVSVAHADSKRKYVYLTRDPTTPLARAAMETAKHTPVN	1800
Db	1741	SAPPDPPKPEYDLELITSCSNVSVAHADSKRKYVYLTRDPTTPLARAAMETAKHTPVN	1800
QY	1801	SWLGNIMYAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPOIQR	1860
Db	1801	SWLGNIMYAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPOIQR	1860
QY	1861	HGLSFAFSLHSYSPGEINRVASCLRLKLGVPPLRVWRHARSVRARLLSQGGRAATCGKYL	1920
Db	1861	HGLSFAFSLHSYSPGEINRVASCLRLKLGVPPLRVWRHARSVRARLLSQGGRAATCGKYL	1920
QY	1921	NWAVRTKLTPIPAASQLDLSWFWAGYSGGDIYHSLSRARPRFWMWCLLLLSVGVGII	1980
Db	1921	NWAVRTKLTPIPAASQLDLSWFWAGYSGGDIYHSLSRARPRFWMWCLLLLSVGVGII	1980
QY	1981	LLPNR 1985	
Db	1981	LLPNR 1985	
RESULT 4			
US-10-029-907-3			
; Sequence 3, Application US/10029907			
; Publication No. US20020142350A1			
; GENERAL INFORMATION:			
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.			
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM			
; TITLE OF INVENTION: HEPATITIS C VIRUS			
; FILE REFERENCE: 13/083			
; CURRENT APPLICATION NUMBER: US/10/029,907			
; CURRENT FILING DATE: 2001-12-21			
; PRIOR APPLICATION NUMBER: 60/257,857			
; PRIOR FILING DATE: 2000-12-22			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 3			
; LENGTH: 2201			
; TYPE: PRT			
; ORGANISM: HCV			
; FEATURES:			
; NAME/KEY: VARIANT			
; LOCATION: 882			
; OTHER INFORMATION: Xaa is Lys or Arg			
; NAME/KEY: VARIANT			
; LOCATION: 1489			
; OTHER INFORMATION: Xaa is Leu			
US-10-029-907-3			
Query Match 99.8%; Score 10442; DB 13; Length 2201;			
Best Local Similarity 99.8%; Pred. No. 0;			
Matches 1981; Conservative 1; Mismatches 3; Indels 0; Gaps 0;			
QY	1	MAPTAYSQOTRGLLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGA	60
Db	217	LAPITAYSQOTRGLLGCIITSLTGRDRNQVEGEVQVSTATQSFATCVNGVCWTVYHGA	276
QY	61	GSKTLAGPKGITOMYTNVDOLVGMQAPPGARSITPCTCGSSDLYLVRHADVIPVRRR	120
Db	277	GSKTLAGPKGITOMYTNVDOLVGMQAPPGARSITPCTCGSSDLYLVRHADVIPVRRR	336
QY	121	GDSRGSLLSPRPVSYLKGSSGGLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM	180
Db	337	GDSRGSLLSPRPVSYLKGSSGGLLCPSGHAGVIFRAAVCTRGVAKAVDFVPVSMETTM	396
QY	181	RSPVFTDNSSPPAVPQTQVAHLHAPTGSGKSTKPAAYAAQGYKVLVLPNSVAATLFGF	240
Db	397	RSPVFTDNSSPPAVPQTQVAHLHAPTGSGKSTKPAAYAAQGYKVLVLPNSVAATLFGF	456
QY	241	AYMSKAHGIDPNIRITGVRTITTTGAPITYSTYKFLADGCGSGAYDIIICDECHSDSTT	300
Db	457	AYMSKAHGIDPNIRITGVRTITTTGAPITYSTYKFLADGCGSGAYDIIICDECHSDSTT	516
QY	301	ILIGICTVLDOAETAGARLVVLATATPGSVTVPHNIEVALSSGEIIPFYKAIPIETI	360
Db	517	ILIGICTVLDOAETAGARLVVLATATPGSVTVPHNIEVALSSGEIIPFYKAIPIETI	576
QY	361	KGGRHLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGT	420
Db	577	KGGRHLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGT	636
QY	421	GFDSVIDNCVTQTVDFSLDPTFTIETTTVPQDAVSRRGRTRGRMGIYFVTPG	480
Db	637	GFDSVIDNCVTQTVDFSLDPTFTIETTTVPQDAVSRRGRTRGRMGIYFVTPG	696
QY	481	ERPSGMFSSVLCCEYDAGCAYELTPAETSVRLRAYLNTPGLPVCQDHLEFWESVFTGL	540
Db	697	ERPSGMFSSVLCCEYDAGCAYELTPAETSVRLRAYLNTPGLPVCQDHLEFWESVFTGL	756
QY	541	THIDAHFLSQTQAGDNFPYLVAQATVCARAQAAPPSPDWQMKCLIRLKPFLHGTPTLL	600
Db	757	THIDAHFLSQTQAGDNFPYLVAQATVCARAQAAPPSPDWQMKCLIRLKPFLHGTPTLL	816
QY	601	YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTSVVIYGR	660
Db	817	YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTSVVIYGR	876
QY	661	IILSGKPAIIPDREVLYREDFEDECASHLYIEQGMQLAEQFKQKATGLIQTATQKQEA	720
Db	877	IILSGKPAIIPDREVLYREDFEDECASHLYIEQGMQLAEQFKQKATGLIQTATQKQEA	936
QY	721	AAPVVESKWRLEAFWAKHWNFIISGLOLAGSLTPCNPAIASLMAFTASITSLTTTQH	780
Db	937	AAPVVESKWRLEAFWAKHWNFIISGLOLAGSLTPCNPAIASLMAFTASITSLTTTQH	996
QY	781	TLLFNILGGVAAQALAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYGAGVAGALVA	840
Db	997	TLLFNILGGVAAQALAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYGAGVAGALVA	1056
QY	841	FKVMSGEMPTEDLVNLLPALISPGALVGVVCAAILRRHVGPGEAVQWNNRLIAFASR	900
Db	1057	FKVMSGEMPTEDLVNLLPALISPGALVGVVCAAILRRHVGPGEAVQWNNRLIAFASR	1116
QY	901	GNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHWINEDCSTPCSGSLRVDWDMIC	960
Db	1117	GNHVSPTHYVPESDAAARVTQILSSLTITQLLKLHWINEDCSTPCSGSLRVDWDMIC	1176
QY	961	TVLTDFTWLOSULLPRLPGVPPFCQRYGVWRGDGMQTTCPGQAQITGHVKNCSMR	1020
Db	1177	TVLTDFTWLOSULLPRLPGVPPFCQRYGVWRGDGMQTTCPGQAQITGHVKNCSMR	1236
QY	1021	IVGPRTCSTNWHGTFPINAYTTGCTSPAPNYSRALWRVAEEYVEVTRVGDHYVTGM	1080
Db	1237	IVGPRTCSTNWHGTFPINAYTTGCTSPAPNYSRALWRVAEEYVEVTRVGDHYVTGM	1296
QY	1081	TTDNVKKCPQVPAPEFFTEVDGVRLEHYAPACKPLLREEVTFVLGNQYLVGSQLPCEPE	1140
Db	1297	TTDNVKKCPQVPAPEFFTEVDGVRLEHYAPACKPLLREEVTFVLGNQYLVGSQLPCEPE	1356
QY	1141	PDVAULTSMLTDPESHITAEAKORLARGSPPSLASSASQLSAPLSKATCTTTRHDSPAD	1200
Db	1357	PDVAULTSMLTDPESHITAEAKORLARGSPPSLASSASQLSAPLSKATCTTTRHDSPAD	1416
QY	1201	LIEANLLWROEMGNITRVESENKVVILDSFEPLQAEDEDESVSPAEIILRRSRKFPFRAM	1260
Db	1417	LIEANLLWROEMGNITRVESENKVVILDSFEPLQAEDEDESVSPAEIILRRSRKFPFRAM	1476
QY	1261	PIWARPDYNPPLLESWKDPDYVPVVGHCPLPPAKAPPIPPRRKRKTVLSESTVSALLA	1320
Db	1477	PIWARPDYNPPLLESWKDPDYVPVVGHCPLPPAKAPPIPPRRKRKTVLSESTVSALLA	1536

```

QY 1321 ELATKTFGSSSESAVDSTATSPDQSDGAGSDVESYSSMPLEGEFGDPLSDGSW 1380
DB 1537 ELATKTFGSSSESAVDSTATSPDQSDGAGSDVESYSSMPLEGEFGDPLSDGSW 1596
QY 1381 STVSEASESDVCCMSYTWGTALITPCAAEETKLPINALSNSLLRHHNLVYATTSSAS 1440
DB 1597 STVSEASESDVCCMSYTWGTALITPCAAEETKLPINALSNSLLRHHNLVYATTSSAS 1656
QY 1441 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVAKALLSVEEAKLTPPHSARKFGYGA 1500
DB 1657 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVAKALLSVEEAKLTPPHSARKFGYGA 1716
QY 1501 DVENLSKAVNHRSVWVKOLLEDETPTDITIMAKNEVFCVQBEKGRKEARLIVPDLG 1560
DB 1717 DVENLSKAVNHRSVWVKOLLEDETPTDITIMAKNEVFCVQBEKGRKEARLIVPDLG 1776
QY 1561 VRYCEKXALYDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAVDTTRCFDS 1620
DB 1777 VRYCEKXALYDVVSTLPOAVMGSSYGFQYSPGQVVEFLVNAWAKKCPMGFAVDTTRCFDS 1836
QY 1621 TVTENDIRVESYQCCDLAPEARQAIRSLTERLIYIGGPLTNSKGQNGYRRCRASGVL 1680
DB 1837 TVTENDIRVESYQCCDLAPEARQAIRSLTERLIYIGGPLTNSKGQNGYRRCRASGVL 1896
QY 1681 TSCGNTLTCLYKAAACRAAKLOCTMLVCGDDLVWICESAGTOEDEASLRAPTEAMTRY 1740
DB 1897 TSCGNTLTCLYKAAACRAAKLOCTMLVCGDDLVWICESAGTOEDEASLRAPTEAMTRY 1956
QY 1741 SAPGDPKPEYDLELITSCSSNVSAVHADSGKVVYLTDPPTPLARAAWETARHTPVN 1800
DB 1957 SAPGDPKPEYDLELITSCSSNVSAVHADSGKVVYLTDPPTPLARAAWETARHTPVN 2016
QY 1801 SWLGNITWAPTLMWARMILMTHFFSILLAOBLEKALDCQIYACYSIEPLDLQIIRL 1860
DB 2017 SWLGNITWAPTLMWARMILMTHFFSILLAOBLEKALDCQIYACYSIEPLDLQIIRL 2076
QY 1861 HGLSAPLSHSYSGEINRVASCLKLVPPPLRVWHRARSVRARLLSQGGRAATCGKLYF 1920
DB 2077 HGLSAPLSHSYSGEINRVASCLKLVPPPLRVWHRARSVRARLLSQGGRAATCGKLYF 2136
QY 1921 NWAVRTKLTPIPAASQLDLSWFWAGYSGGDIYHLSLRARPFWFMCULLLSVGVGIY 1980
DB 2137 NWAVRTKLTPIPAASQLDLSWFWAGYSGGDIYHLSLRARPFWFMCULLLSVGVGIY 2196
QY 1981 LLPNR 1985
DB 2197 LLPNR 2201

```

```

RESULT 5
US-10-309-561-3
; Sequence 3, Application US/10309561
; Publication No. US20030148348A1
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/309,561
; CURRENT FILING DATE: 2002-12-04
; PRIOR APPLICATION NUMBER: US/10/029,907
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 2201
; TYPE: PRT
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 882

```

```

; OTHER INFORMATION: Xaa is Lys or Arg
; FEATURE:
; NAME/KEY: VARIANT
; LOCATION: 1489
; OTHER INFORMATION: Xaa is Leu
US-10-309-561-3

Query Match          99.8%; Score 10442; DB 14; Length 2201;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1981; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 MAPIATYSQOTRGLGCIITISLTGRDRNQVEGVQVSTATQSFATCVNGVCWTVVHGA 60
DB 217 LAPIATYSQOTRGLGCIITISLTGRDRNQVEGVQVSTATQSFATCVNGVCWTVVHGA 276
QY 61 GSKTLGPKGPIITQMTNVQDQDLVQWQAPPGASRLTPTCTGSSDLVLVTHADVIPVRR 120
DB 277 GSKTLGPKGPIITQMTNVQDQDLVQWQAPPGASRLTPTCTGSSDLVLVTHADVIPVRR 336
QY 121 GDSRGSLLSPRVSYLKGSSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDFPVESMETTM 180
DB 337 GDSRGSLLSPRVSYLKGSSGGPLLCPSGHAGVIFRAAVCTRGVAKAVDFPVESMETTM 396
QY 181 RSPVFTDSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLPG 240
DB 397 RSPVFTDSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLPG 456
QY 241 AYSKAGHDIPNRTGVRTITTCAPTITSTYTGFLADGGCGGAYDIIICDECHSTDSTT 300
DB 457 AYSKAGHDIPNRTGVRTITTCAPTITSTYTGFLADGGCGGAYDIIICDECHSTDSTT 516
QY 301 ILGIGTVLDOAETAGARLVLATATPGSVTVPHNPNEEVALSTGTEIPFYGAIPLETI 360
DB 517 ILGIGTVLDOAETAGARLVLATATPGSVTVPHNPNEEVALSTGTEIPFYGAIPLETI 576
QY 361 KGRGHLIFCHSKKCCDELAALKSGLGLNAVAYYRGLDVSIVPTSGDVIVVATDALMTGFT 420
DB 577 KGRGHLIFCHSKKCCDELAALKSGLGLNAVAYYRGLDVSIVPTSGDVIVVATDALMTGFT 636
QY 421 GDFDSVDCNTCTVTQTVDFSLDPTFTTETTTTVPDASVRSORRGRTGRGMGIYRFTVPG 480
DB 637 GDFDSVDCNTCTVTQTVDFSLDPTFTTETTTTVPDASVRSORRGRTGRGMGIYRFTVPG 696
QY 481 ERSFGMEDSVLCECYDAGCAWYELTPAETSVRLRAYLNTPLPVCODHLEFWSVFTGL 540
DB 697 ERSFGMEDSVLCECYDAGCAWYELTPAETSVRLRAYLNTPLPVCODHLEFWSVFTGL 756
QY 541 THIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPSPSDOMWKCLIRLKEPTLHGPTPLL 600
DB 757 THIDAHFLSOTKQAGDNFPYLVAYQATVCARAQAPPSPSDOMWKCLIRLKEPTLHGPTPLL 816
QY 601 YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGSVVLVGR 660
DB 817 YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWVLVGGVLAALAAAYCLTTGSVVLVGR 876
QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKQKAIQLLTATKQAEA 720
DB 877 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKQKAIQLLTATKQAEA 936
QY 721 AAPVVEKWTLEAFWAKHMMNFSGIOYLAGLSTLPGNPAIASLMAFTASITSPLOTQH 780
DB 937 AAPVVEKWTLEAFWAKHMMNFSGIOYLAGLSTLPGNPAIASLMAFTASITSPLOTQH 996
QY 781 TLLFNILGGWVAQAOLAPPASAASAFVAGIAGAAVGSIGLKVLDVILLAGYAGVAGALVA 840
DB 997 TLLFNILGGWVAQAOLAPPASAASAFVAGIAGAAVGSIGLKVLDVILLAGYAGVAGALVA 1056
QY 841 FKVMGEMPTSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVQVWNNRLIAFASR 900
DB 1057 FKVMGEMPTSTEDLVNLLPAILSPCALVGVVCAAILRRHVGPGEVQVWNNRLIAFASR 1116
QY 901 GNHVSPTHYVDESAAAARVTOILSSLTITOLLKRIHOWINEDCSTPCSGSWLRDYMWDIC 960

```

Db 1117 GNHVSPTHVVPESDAAARVTQILSSLTITQLLKRLLHQMINEDCSTPCSSWLRDVMWIC 1176  
Qy 961 TVLTDKFWLQSKLLPRLPGVPPFFCQCGYKGVWGRGDGIMQITPCGQAQITGHVXGSMR 1020  
Db 1177 TVLTDKFWLQSKLLPRLPGVPPFFCQCGYKGVWGRGDGIMQITPCGQAQITGHVXGSMR 1236  
Qy 1021 IVGPRTCSNTHGTPPINAYTTGCTPSPAPNYSRMLRWAAEEVEVTRVGDHRYVTGM 1080  
Db 1237 IVGPRTCSNTHGTPPINAYTTGCTPSPAPNYSRMLRWAAEEVEVTRVGDHRYVTGM 1296  
Qy 1081 TTDNVKFCQVPAPEFFTEVDGVRLLHRYAPACKPLLRBEVTFVLGYNQVLVGSQPCPE 1140  
Db 1297 TTDNVKFCQVPAPEFFTEVDGVRLLHRYAPACKPLLRBEVTFVLGYNQVLVGSQPCPE 1356  
Qy 1141 PDVAULTSMLTDPHSITATKARRLARGSPSLASSASQLSAPLSKATCTTTRHDSPAD 1200  
Db 1357 PDVAULTSMLTDPHSITATKARRLARGSPSLASSASQLSAPLSKATCTTTRHDSPAD 1416  
Qy 1201 LIEANLLWRQBMGNITRVESENKVVILDSPEPLQAEDEREVSPAEILRRSRKFPBRAM 1260  
Db 1417 LIEANLLWRQBMGNITRVESENKVVILDSPEPLQAEDEREVSPAEILRRSRKFPBRAM 1476  
Qy 1261 PIWARPDPNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1320  
Db 1477 PIWARPDPNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSSALA 1536  
Qy 1321 ELATKTFGSSSESSAVDSGTATASPDQSDGDSVESYSMPLEGEPPDLSDGSM 1380  
Db 1537 ELATKTFGSSSESSAVDSGTATASPDQSDGDSVESYSMPLEGEPPDLSDGSM 1596  
Qy 1381 STVSEEAESDVVCCSMSTWTGALITPCAARETKLPINALNSLLRHNLVYATTSRAS 1440  
Db 1597 STVSEEAESDVVCCSMSTWTGALITPCAARETKLPINALNSLLRHNLVYATTSRAS 1656  
Qy 1441 LRQKVTDFRLQVLDHYRDVLKEMKAKASTVKALLSVEBACKLTPPHSARSKFYGAK 1500  
Db 1657 LRQKVTDFRLQVLDHYRDVLKEMKAKASTVKALLSVEBACKLTPPHSARSKFYGAK 1716  
Qy 1501 DVRLNSSKANVHRSWKDLEDTEPTDITIMAKNEVFCVQPKGGRKPARLIVFPDLG 1560  
Db 1717 DVRLNSSKANVHRSWKDLEDTEPTDITIMAKNEVFCVQPKGGRKPARLIVFPDLG 1776  
Qy 1561 VRCEKVALYDVSTLPOAVMGSSVGFQYSPQGVFELVNAWKAACKCPMGFAYTRCFDS 1620  
Db 1777 VRCEKVALYDVSTLPOAVMGSSVGFQYSPQGVFELVNAWKAACKCPMGFAYTRCFDS 1836  
Qy 1621 TVTENDIRVEESIYQCCDLAPAPARQAIRSLTERLYIGGLTNSKGQNGYRRCRASGVL 1680  
Db 1837 TVTENDIRVEESIYQCCDLAPAPARQAIRSLTERLYIGGLTNSKGQNGYRRCRASGVL 1896  
Qy 1681 TSCGNTLTCVYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTOEDASLPAFTAMTRY 1740  
Db 1897 TSCGNTLTCVYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTOEDASLPAFTAMTRY 1956  
Qy 1741 SAPPGDPKPEYDLELITSCSSNVSVHADSGKRVYVLTDRPTTPLARAAMETARHTPVN 1800  
Db 1957 SAPPGDPKPEYDLELITSCSSNVSVHADSGKRVYVLTDRPTTPLARAAMETARHTPVN 2016  
Qy 1801 SWLGNIIMYAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIQL 1860  
Db 2017 SWLGNIIMYAPTLWARMILMTHFFSILLAQOLEKALDCQIYGACYSIEPLDLPQIQL 2076  
Qy 1861 HGLSFAFSLHSVSGEINRVASCLKLGVPPLVRWHRARSVRARLLSQGGRAATCGKYL 1920  
Db 2077 HGLSFAFSLHSVSGEINRVASCLKLGVPPLVRWHRARSVRARLLSQGGRAATCGKYL 2136  
Qy 1921 NWAVRTKLLTPIPAASQDLSSVFAVAGSGGDIYHSLSRARPFWMCULLLSVGVGIY 1980  
Db 2137 NWAVRTKLLTPIPAASQDLSSVFAVAGSGGDIYHSLSRARPFWMCULLLSVGVGIY 2196  
Qy 1981 LLPNR 1985  
Db 2197 LLPNR 2201

RESULT 6  
US-10-789-355-3  
; Sequence 3, Application US/10789355  
; Publication No. US20040180333A1  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/789,355  
; CURRENT FILING DATE: 2004-02-27  
; PRIOR APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 2201  
; TYPE: PRT  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 882  
; OTHER INFORMATION: Xaa is Lys or Arg  
; FEATURE:  
; NAME/KEY: VARIANT  
; LOCATION: 1489  
; OTHER INFORMATION: Xaa is Leu  
US-10-789-355-3

Query Match 99.8%; Score 10442; DB 16; Length 2201;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 1981; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MAPITAYSQQTGRLGCIITSLTGRDRNOVEGEVQVYVSTATQSFATCNGVCVTVYHGA 60  
Db 217 LAPITAYSQQTGRLGCIITSLTGRDRNOVEGEVQVYVSTATQSFATCNGVCVTVYHGA 276  
Qy 61 GSKTLAGPKGPIITOMYTNVDQDLVGVQAPPCARSLTCTCGSSDLYLVTTHADVIVRRR 120  
Db 277 GSKTLAGPKGPIITOMYTNVDQDLVGVQAPPCARSLTCTCGSSDLYLVTTHADVIVRRR 336  
Qy 121 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
Db 337 GDSRGSLLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 396  
Qy 181 RSPVFTDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLFG 240  
Db 397 RSPVFTDNSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLFG 456  
Qy 241 AYMSKAHGDIPNRTGVRTITTCGAPITYSYGKFLADGGCSGGAYDIIICDECHSTDTT 300  
Db 457 AYMSKAHGDIPNRTGVRTITTCGAPITYSYGKFLADGGCSGGAYDIIICDECHSTDTT 516  
Qy 301 ILGIGTVLDOAETAGARLVVLTATPPGVTVPHPNIEEVALSSTGEIPFGKAIPIETI 360  
Db 517 ILGIGTVLDOAETAGARLVVLTATPPGVTVPHPNIEEVALSSTGEIPFGKAIPIETI 576  
Qy 361 KGRHILFCHSKKKKDELAALKSLGLNAVAYVRGLDVSVIPTSQGVIVVATDALMTGFT 420  
Db 577 KGRHILFCHSKKKKDELAALKSLGLNAVAYVRGLDVSVIPTSQGVIVVATDALMTGFT 636  
Qy 421 GDFDSVIDCNTCTVQTVDFSLDPTFTIETTVPQDAVRSQRRGRTGRGMGIYRFVTPG 480  
Db 637 GDFDSVIDCNTCTVQTVDFSLDPTFTIETTVPQDAVRSQRRGRTGRGMGIYRFVTPG 696  
Qy 481 ERPSGMFDDSVLCECYDAGCAWVELTPAETSVELRAYLNTPGLPVCDHLEFWESVFTGL 540  
Db 697 ERPSGMFDDSVLCECYDAGCAWVELTPAETSVELRAYLNTPGLPVCDHLEFWESVFTGL 756





Db 337 GDSRGLSPRPVSYLKGSGCLLCPGSHAVGIFRAAVCTRGVAKAYDFVPVSMETTM 396  
Qy 181 RSPVPTDNSSPAVDQTTQVQVHLHAPTQSGSKTKVPAAYAAQGYKVLVNLNPSVAATLFGF 240  
Db 397 RSPVPTDNSSPAVDQTTQVQVHLHAPTQSGSKTKVPAAYAAQGYKVLVNLNPSVAATLFGF 456  
Qy 241 AYMSKAHGIDNIRGVRTITIGAPITYSTYKGLADGCGSGGAYDIIICDECHSTDSTT 300  
Db 457 AYMSKAHGIDNIRGVRTITIGAPITYSTYKGLADGCGSGGAYDIIICDECHSTDSTT 516  
Qy 301 ILGIGTVLDQAEAGARLWLATATPPGSSVTPPHNIEEVALSSSTGEIPFYKGAIPDIETI 360  
Db 517 ILGIGTVLDQAEAGARLWLATATPPGSSVTPPHNIEEVALSSSTGEIPFYKGAIPDIETI 576  
Qy 361 KGRHLLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIVVATDALMTGFT 420  
Db 577 KGRHLLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIVVATDALMTGFT 636  
Qy 421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQAVRSORRGRTGRMGIVRFVTPG 480  
Db 637 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQAVRSORRGRTGRMGIVRFVTPG 696  
Qy 481 ERPSGMFSSVLCBDCYDAGCWAYBELTPAETSVRLRAYLNTFGLPVCQDHLFEFWSVFTGL 540  
Db 697 ERPSGMFSSVLCBDCYDAGCWAYBELTPAETSVRLRAYLNTFGLPVCQDHLFEFWSVFTGL 756  
Qy 541 THIDAHFSLQTKQAGDNFPYLVAYQATVCARQAAPPSPNDQWKCLIRLKTILHGPPL 600  
Db 757 THIDAHFSLQTKQAGDNFPYLVAYQATVCARQAAPPSPNDQWKCLIRLKTILHGPPL 816  
Qy 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTVTWLVGGVLAALAAAYCLTTGSSVVIIGR 660  
Db 817 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTVTWLVGGVLAALAAAYCLTTGSSVVIIGR 876  
Qy 661 IILSGKPAIIPDRVLRYREFDEMEECASHLPYIBQGMQLARQFKQKAIQLLOTATKQAEA 720  
Db 877 IILSGKPAIIPDRVLRYREFDEMEECASHLPYIBQGMQLARQFKQKAIQLLOTATKQAEA 936  
Qy 721 AAPVVESKWRTEAFWAKHWNFIISGQYLAGLSTLPGNPAIASIMAPTASITSPLTQ 780  
Db 937 AAPVVESKWRTEAFWAKHWNFIISGQYLAGLSTLPGNPAIASIMAPTASITSPLTQ 996  
Qy 781 TLLPNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLGVLDVILLAGYAGAGALVA 840  
Db 997 TLLPNILGGWVAQALAPPSAASAFVAGIAGAAVGSIGLGVLDVILLAGYAGAGALVA 1056  
Qy 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 900  
Db 1057 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 1116  
Qy 901 GNVHSPHYYPEDSAAARVTQILSSLIITQLKRLHOWINEDCSTPCSGWLRDWDWIC 960  
Db 1117 GNVHSPHYYPEDSAAARVTQILSSLIITQLKRLHOWINEDCSTPCSGWLRDWDWIC 1176  
Qy 961 TVLTDFTWLTQSKLLPRLPGVPFPSCORGKGVWRGDIQMTTCCPQCAQITGHVKNCSMR 1020  
Db 1177 TVLTDFTWLTQSKLLPRLPGVPFPSCORGKGVWRGDIQMTTCCPQCAQITGHVKNCSMR 1236  
Qy 1021 IVGERTCSNTWHGTFPINAVTTGCTPSPAPNYSRALWRVAAEYVEVTRVGDPHYVTGM 1080  
Db 1237 IVGERTCSNTWHGTFPINAVTTGCTPSPAPNYSRALWRVAAEYVEVTRVGDPHYVTGM 1296  
Qy 1081 TTDNVKCPQVPAPEFEFTEVDGVLRLHYAPACKPDLLEEVTVFLVGLNQYLVGSLPCEPE 1140  
Db 1297 TTDNVKCPQVPAPEFEFTEVDGVLRLHYAPACKPDLLEEVTVFLVGLNQYLVGSLPCEPE 1356  
Qy 1141 PDVAVLTSMLTDPDSHITAETAKRLARGSPSPASSASQLSAPSLKATCTTTRHDSDDAD 1200  
Db 1357 PDVAVLTSMLTDPDSHITAETAKRLARGSPSPASSASQLSAPSLKATCTTTRHDSDDAD 1416  
Qy 1201 LIEANLLWRQEMGGNITRVSEENKVILDSFEPLQAEDEDEREVSVPABEILRRSKFFPRAM 1260  
Db 1417 LIEANLLWRQEMGGNITRVSEENKVILDSFEPLQAEDEDEREVSVPABEILRRSKFFPRAM 1476

Qy 1261 PIWARPDPNPLLESWSKDDPDYVPPVHGCPLPPAKAPIPPRRKRRTTVLSESTVSSAIA 1320  
Db 1477 PIWARPDPNPLLESWSKDDPDYVPPVHGCPLPPAKAPIPPRRKRRTTVLSESTVSSAIA 1536  
Qy 1321 ELATKTFGSSSESSAVDSGTATASPDQSDDDGSDVESYSMPPLLEGEPDPLSDGSW 1380  
Db 1537 ELATKTFGSSSESSAVDSGTATASPDQSDDDGSDVESYSMPPLLEGEPDPLSDGSW 1596  
Qy 1381 STVSEERASEDVVCCSMSTYWTGALITPCAAEBETKLPIINALSNSLLRHHNLVATTSSRAS 1440  
Db 1597 STVSEERASEDVVCCSMSTYWTGALITPCAAEBETKLPIINALSNSLLRHHNLVATTSSRAS 1656  
Qy 1441 LRQKVTFRDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEBACLTPPHSARSFEGYAK 1500  
Db 1657 LRQKVTFRDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEBACLTPPHSARSFEGYAK 1716  
Qy 1501 DVRLNSSKAVNHIRSVMKDLLEDTEPIDTITIMAKNEVFCVQPEKGGKPKARLIVFPDLG 1560  
Db 1717 DVRLNSSKAVNHIRSVMKDLLEDTEPIDTITIMAKNEVFCVQPEKGGKPKARLIVFPDLG 1776  
Qy 1561 VRVCEKALYDVVSTLPOAVMGSSYGFQYSPGQVFEFLVNAWKAKCPCMGFAYDTRCFDS 1620  
Db 1777 VRVCEKALYDVVSTLPOAVMGSSYGFQYSPGQVFEFLVNAWKAKCPCMGFAYDTRCFDS 1836  
Qy 1621 TVTENDIRVBESIIYCCDLAPARQAIIRSLTERLYIGGPLTNSKGONCGYRRCRASGVLT 1680  
Db 1837 TVTENDIRVBESIIYCCDLAPARQAIIRSLTERLYIGGPLTNSKGONCGYRRCRASGVLT 1896  
Qy 1681 TSCGNTITCYLKAACRAAKLODCTMLVCGDDLVAVICESAGTOEDRASIAFTEAMTRY 1740  
Db 1897 TSCGNTITCYLKAACRAAKLODCTMLVCGDDLVAVICESAGTOEDRASIAFTEAMTRY 1956  
Qy 1741 SAPPDGPPEYDOLELITSCSSNVSVVAHDASGKRVYVLTDRDPTPLARAAMETARHTPEVN 1800  
Db 1957 SAPPDGPPEYDOLELITSCSSNVSVVAHDASGKRVYVLTDRDPTPLARAAMETARHTPEVN 2016  
Qy 1801 SWLGNIIIMYAPTLMARMILMTHFFSILLAOEKLKALDCQIYGACYSIEPLDLPQIIQRL 1860  
Db 2017 SWLGNIIIMYAPTLMARMILMTHFFSILLAOEKLKALDCQIYGACYSIEPLDLPQIIQRL 2076  
Qy 1861 HGLSAPLSHSYSGEINRVASCLRLKGLVPLRVWRHARSVRARLLSQGGRAATCGKYL 1920  
Db 2077 HGLSAPLSHSYSGEINRVASCLRLKGLVPLRVWRHARSVRARLLSQGGRAATCGKYL 2136  
Qy 1921 NWAVRTKLTPTPAASQDLSSWFAVAGYSGGDIYHLSLRARPRFWMCLLLSVGVGIY 1980  
Db 2137 NWAVRTKLTPTPAASQDLSSWFAVAGYSGGDIYHLSLRARPRFWMCLLLSVGVGIY 2196  
Qy 1981 LLPNR 1985  
Db 2197 LLPNR 2201

## RESULT 8

US-10-085-476-2  
; Sequence 2, Application US/10085476  
; Publication No. US20020164722A1  
; GENERAL INFORMATION:  
; APPLICANT: De Francesco, Raffaele  
; APPLICANT: Tomei, Licia  
; APPLICANT: Behrens, Sven-Brik  
; TITLE OF INVENTION: METHOD FOR REPRODUCING IN VITRO THE  
; TITLE OF INVENTION: RNA-DEPENDENT RNA POLYMERASE AND TERMINAL NUCLEOTIDYL  
; TITLE OF INVENTION: TRANSFERASE ACTIVITIES ENCODED BY HEPATITIS C VIRUS (HCV)  
; FILE REFERENCE: IT0002PCA  
; CURRENT APPLICATION NUMBER: US/10/085,476  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: 08/952,981  
; PRIOR FILING DATE: 1998-03-23  
; PRIOR APPLICATION NUMBER: PCT/IT96/00106  
; PRIOR FILING DATE: 1996-05-24  
; PRIOR APPLICATION NUMBER: RM95A000343

; PRIOR FILLING DATE: 1995-05-25 ; NUMBER OF SEQ ID NOS: 14 ; SOFTWARE: FastSeq for Windows Version 4.0 ; SEQ ID NO 2 ; LENGTH: 2201 ; TYPE: PRT ; ORGANISM: cDNA clone pCD (38-9.4) ; US-10-085-476-2									
Query Match 97.2%; Score 10170; DB 13; Length 2201; Best Local Similarity 96.7%; Pred. No. 0; Matches 1919; Conservative 36; Mismatches 30; Indels 0; Gaps 0;									
QY	1	MAPITAYSOOTRGLLCIITSLTRDRNQVEGVQVSTATQSFATCNGVCWTVYHGA	60						
DB	217	LAPITAYSOOTRGLLCIITSLTRDRNQVEGVQVSTATQSFATCNGVCWTVYHGA	276						
QY	61	GSKTAGPKGPIQMTYTNVDQDLVGMQAPGARSLLPTCGSSDLYLVTRHADVIPIVRRR	120						
DB	277	GSKTLAAPKGPITQMTYTNVDQDLVGMQAPGARSLLPTCGSSDLYLVTRHADVIPIVRRR	336						
QY	121	GSRGSLSPRPVSYLKGSGGELLCPSPGHAVGIFRAAVCTRGVAKAVDFPVPVSMETTM	180						
DB	337	GSRGSLSPRPVSYLKGSGGELLCPSPGHAVGIFRAAVCTRGVAKAVDFPVPVSMETTM	396						
QY	181	RSVPFTDNSPPRAVQTFQVAHLHAPTGSCKSTKVPAAVAAQGYKVLVNLNPSVAATLFGF	240						
DB	397	RSVPFTDNSPPRAVQTFQVAHLHAPTGSCKSTKVPAAVAAQGYKVLVNLNPSVAATLFGF	456						
QY	241	AYMSKAHGIDPNIRTGRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT	300						
DB	457	AYMSKAHGIDPNIRTGRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT	516						
QY	301	ILGIGTVLDOAETAGARLVVLAATATPPGSVTVPHPNIEEVALSGTGEIPYKAIPIETI	360						
DB	517	ILGIGTVLDOAETAGARLVVLAATATPPGSVTVPHPNIEEVALSGTGEIPYKAIPIETI	576						
QY	361	XGRGHLIFCHSKKKDELAAKLSGLGINAVAYYRGDVSIVPTSGDVIIVATDALMTGFT	420						
DB	577	XGRGHLIFCHSKKKDELAAKLSGLGINAVAYYRGDVSIVPTSGDVIIVATDALMTGFT	636						
QY	421	GFDSVIDCNTCTQTVDFSLDFTFTTETTVQDAVRSQRGRGRGRGMYRFRVTPG	480						
DB	637	GFDSVIDCNTCTQTVDFSLDFTFTTETTVQDAVRSQRGRGRGRGMYRFRVTPG	696						
QY	481	ERPSGMFDSVLCCEYDAGCAWVELTAPETSRLRAYLNTPLPVCDHLEFWSVPTGL	540						
DB	697	ERPSGMFDSVLCCEYDAGCAWVELTAPETSRLRAYLNTPLPVCDHLEFWSVPTGL	756						
QY	541	THIDAHFLSOTKAGDNFPVLVAYQATVCARAQAPPSDMMKCLTLKPTLHGPTLL	600						
DB	757	THIDAHFLSOTKAGDNFPVLVAYQATVCARAQAPPSDMMKCLTLKPTLHGPTLL	816						
QY	601	YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTTSTVLVGVGLAALAAAYCLITGSVVIIVGR	660						
DB	817	YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTTSTVLVGVGLAALAAAYCLITGSVVIIVGR	876						
QY	661	IILSGKPAIIPDREVLYRBPDEMBECASHLPYIEQGMQLAEQFKQKAILLOTATQKAE	720						
DB	877	IILSGKPAIIPDREVLYRBPDEMBECASHLPYIEQGMQLAEQFKQKAILLOTATQKAE	936						
QY	721	AAPVSKWKTLEAFWAKHWNFTISGTYLAGLSTLPGNPAIASLMAFTASITSLPTQH	780						
DB	937	AAPVSKWKTLEAFWAKHWNFTISGTYLAGLSTLPGNPAIASLMAFTASITSLPTQH	996						
QY	781	TLLENILGGVAAQAAPPASAFVAGIAGAAVGSIGLKVLDIILAGYAGAGVALVA	840						
DB	997	TLLENILGGVAAQAAPPASAFVAGIAGAAVGSIGLKVLDIILAGYAGAGVALVA	1056						
QY	841	FKVMSGMPSTEDLWNLPAILSPGALVGVVCAAILRRHVGPGEVQWNNRLIAFASR	900						
DB	1057	FKVMSGMPSTEDLWNLPAILSPGALVGVVCAAILRRHVGPGEVQWNNRLIAFASR	1116						



QY 1741 SAPPDPPKPYDLELITSCSSNVSAHDASGRVYVLTREDPTTPLARAWEHTARHTPVN 1800  
DB 2766 SAPPDPPKPYDLELITSCSSNVSAHDASGRVYVLTREDPTTPLARAWEHTARHTPVN 2825  
QY 1801 SWLGNIMYAPTLMARMILMTWTFSSILLAQOEKALDCQIYGACYSIEPLDLPQIIQRL 1860  
DB 2826 SWLGNIMYAPTLMARMILMTWTFSSILLAQOEKALDCQIYGACYSIEPLDLPQIIQRL 2885  
QY 1861 HGLSAPSLHSYSGEINRVASCLRLKLGVPPLRVHRARSVRARLLSQSGRAATCGKYL 1920  
DB 2886 HGLSAPSLHSYSGEINRVASCLRLKLGVPPLRVHRARSVRARLLSQSGRAATCGKYL 2945  
QY 1921 NNAVRKLLKLTPIPAASQDLDSWVFAGYSGGDIYHLSGRPRWFWMCLLLSVGVGY 1980  
DB 2946 NNAVRKLLKLTPIPAASQDLDSWVFAGYSGGDIYHLSGRPRWFWMCLLLSVGVGY 3005  
QY 1981 LILPNR 1985  
DB 3006 LILPNR 3010  
RESULT 10  
US-10-259-275-40  
; Sequence 40, Application US/10259275  
; Publication No. US20030125541A1  
; GENERAL INFORMATION:  
; APPLICANT: Lemon, Stanley M.  
; TITLE OF INVENTION: REPLICATION COMPETENT HEPATITIS C VIRUS AND METHODS OF USE  
; FILE REFERENCE: 265.0007 0120  
; CURRENT APPLICATION NUMBER: US/10/259,275  
; CURRENT FILING DATE: 2003-01-13  
; PRIORITY APPLICATION NUMBER: US 60/171,909  
; PRIORITY FILING DATE: 1999-12-23  
; PRIOR APPLICATION NUMBER: US 09/747,419  
; PRIOR FILING DATE: 2000-12-23  
; PRIOR APPLICATION NUMBER: US 60/325,236  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: US 60/338,123  
; PRIOR FILING DATE: 2001-11-13  
; NUMBER OF SEQ IDS: 73  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 40  
; LENGTH: 2985  
; TYPE: PRT  
; ORGANISM: ARTIFICIAL  
; FEATURE:  
; OTHER INFORMATION: amino acid sequence encoded by nucleotides 2077-11121 of SEQ ID  
; OTHER INFORMATION: NO:39  
US-10-259-275-40  
Query Match 92.8%; Score 9710; DB 14; Length 2985;  
Best Local Similarity 92.9%; Pred. No. 0;  
Matches 1831; Conservative 59; Mismatches 66; Indels 14; Gaps 3;  
QY 25 RDRNOVEGEVQVSTATQSFATCVGCVTVVHGAGSKTLGPKGITQYNTNVDDLV 84  
DB 1021 RDNQVEGEVQVSTATQSFATCVGCVTVVHGAGSKTLGPKGITQYNTNVDDLV 1080  
QY 85 GWQAPPQARSITPTCGSSDLYLVTRHADVIPVRRRGRSGSLSPRVSYLKGSSGGL 144  
DB 1081 GWPAPPQARSITPTCGSSDLYLVTRHADVIPVRRRGRSGSLSPRVSYLKGSSGGL 1140  
QY 145 LCPSGHVGIFRAAVCTRGVAKAVDFVPVSMETMTMSPVFTDNSSPPAVPQTFQVAHLH 204  
DB 1141 LCPSGHVGIFRAAVCTRGVAKAVDFVPVSMETMTMSPVFTDNSSPPAVPQTFQVAHLH 1200  
QY 205 APTGSGKSTKVPAAAYAAQYKVLNPSVAATLGFGRYMSKAHGDIDNIRTVGRTITGA 264  
DB 1201 APTGSGKSTKVPAAAYAAQYKVLNPSVAATLGFGRYMSKAHGDIDNIRTVGRTITGA 1260  
QY 265 PITVSTYTKFLADGGCGSGGAYDIIICDECHSTDSTTILGIGTVLDQAGTARLVVLATA 324

DB 1261 PITVSTYTKFLADGGCGSGGAYDIIICDECHSTDSTTILGIGTVLDQAGTARLVVLATA 1320  
QY 325 TTPGGSVTPHPNIEEVALSSTGEIIPYKKAIPETIKGGRHLIFCHSKKCCDELAALSG 384  
DB 1321 TTPGGSVTPHPNIEEVALSSTGEIIPYKKAIPETIKGGRHLIFCHSKKCCDELAALSG 1380  
QY 385 LGINAVAYYRGDLDSVIPSPTSGDVIIVATDALTGTGDFSDSVIDCMTCVTQTVDVDSLDP 444  
DB 1381 LGINAVAYYRGDLDSVIPSPTSGDVIIVATDALTGTGDFSDSVIDCMTCVTQTVDVDSLDP 1440  
QY 445 FTIETTTPQDVAVSQRGRGTGRGMGIYRVFTPGRSPGMPDSSVLCRCYDAGCAWYE 504  
DB 1441 FTIETTTPQDVAVSQRGRGTGRGMGIYRVFTPGRSPGMPDSSVLCRCYDAGCAWYE 1500  
QY 505 LTPAETSIVLRAYLNTPLGVCODHLEFWESVFTGLTHIDAHFLSOTKQAGDNFFYLVA 564  
DB 1501 LTPAETSIVLRAYLNTPLGVCODHLEFWESVFTGLTHIDAHFLSOTKQAGDNFFYLVA 1560  
QY 565 QATVCARAOAPPPSDQMWKCLIRLXPTLHGPTPLLYRLGAVQNEVTTTTHTPKYIMACM 624  
DB 1561 QATVCARAOAPPPSDQMWKCLIRLXPTLHGPTPLLYRLGAVQNEVTTTTHTPKYIMACM 1620  
QY 625 SADLEVVTSTWLVGCVLAALAAAYCLTTGTSVVIIVGRIILSGKPAIIPDRVILVREDDME 684  
DB 1621 SADLEVVTSTWLVGCVLAALAAAYCLTTGTSVVIIVGRIILSGRPAVVPDRVILVREDDME 1680  
QY 685 ECASHLPYIEQGMOLAEQFKQKALGLQATQKAEAAAAPVVEKSKWTLAEAFNAKHMNFI 744  
DB 1681 ECASHLPYIEQGMOLAEQFKQKALGLQATQKAEAAAAPVVEKSKWTLAEAFNAKHMNFI 1740  
QY 745 SGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOHTLFLNILGWVAAQLAPPSAASAF 804  
DB 1741 SGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOHTLFLNILGWVAAQLAPPSAASAF 1800  
QY 805 VGAGIAGAAVSGISGLKVLVDILAGVAGVAGALVAFKVMGSGMPSTEDLVNLLPAILSP 864  
DB 1801 VGAGIAGAAVSGISGLKVLVDILAGVAGVAGALVAFKVMGSGMPSTEDLVNLLPAILSP 1860  
QY 865 GALVVGVCAAIILRRHVGPGEVAVQMNRLIAFASRGNHVSPTHYVYPESDAAARVTOILS 924  
DB 1861 GALVVGVCAAIILRRHVGPGEVAVQMNRLIAFASRGNHVSPTHYVYPESDAAARVTOILS 1920  
QY 925 SLTITQLLKLHOWINEDCSTPCSGSLRDVDMICTVLDTDKTLWLSKLLPRLPGVPPF 984  
DB 1921 SLTITQLLKLHOWINEDCSTPCSGSLRDVDMICTVLDTDKTLWLSKLLPRLPGVPPF 1980  
QY 985 SCQRYKGVWRGGINQWTTCPGCAQITGHVKGSMRIIVGPRCTSNTHGTFPINAYTTGP 1044  
DB 1981 SCQRYKGVWRGGINQWTTCPGCAQITGHVKGSMRIIVGPRCTSNTHGTFPINAYTTGP 2040  
QY 1045 CTSPAPNYSRALWRVAABEYVETRVGDFHYVTGTTDNVCKPCQVPAPEFTEVDGVR 1104  
DB 2041 CTSPAPNYSRALWRVAABEYVETRVGDFHYVTGTTDNVCKPCQVPAPEFTEVDGVR 2100  
QY 1105 LHRYPACKPLREEVTFVLGNQLVGSQJL-----PCEPEPDAVLTSMLTDPSSHATAE 1159  
DB 2101 LHRYPACKPLREEVTFVLGNQLVGSQJL-----PCEPEPDAVLTSMLTDPSSHATAE 2155  
QY 1160 TAKRRILARGPPSLASSASQLSAPSLKATCTTTH-----DSPADLIEANLLWQEMGN 1215  
DB 2156 LVGSQARGPPSLASSASQLSAPSLKATCTTTHSSYNLSDSPVDLIAALLWQEMGN 2215  
QY 1216 ITRVESNKKVVLDSPEPLQAEEDEREVSVAELILRRSRKFFRAMPPIWARPDYNNPILLES 1275  
DB 2216 ITRVESNKKVVLDSPEPLQAEEDEREVSVAELILRRSRKFFRAMPPIWARPDYNNPILLES 2275  
QY 1276 WKPDYVYPPVHGCPLFPAPAKPIPPPRKRTVVLSESTVSSALAEATKTFGSESASV 1335  
DB 2276 WKPDYVYPPVHGCPLFPAPAKPIPPPRKRTVVLSESTVSSALAEATKTFGSESASV 2335  
QY 1336 DSGTATASPOPSDDGADGSDVSESSMPPLEGPGDPLSDGSMSTVSEASDVVCCS 1395

Db 2336 DSGTATAPDQTSDDGKSDAESCSPPELGEPCDPLSDGSMSTVEEAGEVSVCCS 2395

QY 1396 NSYTWGTGALITPCAAEETKLPINALSNSLLRHNLVYATTSASLRQKKVTFDRLOQLD 1455

Db 2396 NSYTWGTGALITPCAAEESKLPINALSNSLLRHNLVYATTSASLRQKKVTFDRLOQLD 2455

QY 1456 DHYRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFGYGAQDVRLNSKAVNHHS 1515

Db 2456 DHYRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSKFGYGAQDVRLNSKAVNHHS 2515

QY 1516 VKOLLEDETETIDITIMAKNEVFCVQPEKGGKPARLIVFPDLGVRCCKMALYDVVST 1575

Db 2516 VWEDLLEDTVPIDITIMAKNEVFCVQPEKGGKPARLIVFPDLGVRCCKMALYDVVST 2575

QY 1576 LPOAVMGSSYGFQSPGQRFVFLVNAWKAKCPMGFAVDTRCFDSTVTENDIRVEESIYQ 1635

Db 2576 LPOAVMGSSYGFQSPGQRFVFLVNAWKAKCPMGFAVDTRCFDSTVTENDIRVEESIYQ 2635

QY 1636 CCDLAPEARQAIRSLTERLYIGGPLTNSKGQNGCYRRRCRAGSVLTSCGNTLTCYLKASA 1695

Db 2636 CCDLAPEARQAKSLTERLYIGGPLTNSKGQNGCYRRRCRAGSVLTSCGNTLTCYLKASA 2695

QY 1696 ACRAAKLOCTMVLGCDLWVICSAGTQEDBEASLRAFTTEAMTRYSAAPGDPKPEYDLE 1755

Db 2696 ACRAAKLOCTMVLGCDLWVICSAGTQEDBEASLRAFTTEAMTRYSAAPGDPKPEYDLE 1815

QY 1756 LITSCSSNVSAHDASGKRVYLTDPPTPLARAAWETARHTPVNSWLNIIIMVAPTUWA 1815

Db 2756 LITSCSSNVSAHDASGKRVYLTDPPTPLARAAWETARHTPVNSWLNIIIMVAPTUWA 2815

QY 1816 RMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORLHGLSAFSLHSYSPGE 1875

Db 2816 RMILMTHFFSILLAOEQLKALDCQIYGACYSIEPLDLPQIIORLHGLSAFSLHSYSPGE 2875

QY 1876 INRVASCLRKUGVPLRVRHRSVRRLKSQGGRAATCGKILFNWAVRTKLITPIPA 1935

Db 2876 INRVASCLRKUGVPLRVRHRSVRRLKSQGGRAATCGKILFNWAVRTKLITPIPA 2935

QY 1936 ASQDLASWFWAGYSGGDIYHLSRARPFWMLCLLLSVGVGIYLLPNR 1985

Db 2936 ASQDLASWFWAGYSGGDIYHLSRARPFWMLCLLLSVGVGIYLLPNR 2985

RESULT 11

US-10-296-734-406

; Sequence 406, Application US/10296734

; Publication No. US20040054137A1

; GENERAL INFORMATION:

; APPLICANT: Thompson, Scott A

; TITLE OF INVENTION: Synthetic molecules and uses therefor

; FILE REFERENCE: Savine

; CURRENT APPLICATION NUMBER: US/10/296,734

; CURRENT FILING DATE: 2003-08-04

; PRIOR APPLICATION NUMBER: AU P07761/00

; PRIOR FILING DATE: 2000-05-26

; NUMBER OF SEQ ID NOS: 1507

; SOFTWARE: PatentIn version 3.2

; SEQ ID NO 406

; LENGTH: 3011

; TYPE: PRT

; ORGANISM: Artificial

; FEATURE:

; OTHER INFORMATION: HepC la consensus polyprotein

US-10-296-734-406

Query Match 90.3%; Score 9454.5; DB 15; Length 3011;

Best Local Similarity 88.0%; Pred. No. 0;

Matches 1749; Conservative 134; Mismatches 101; Indels 3; Gaps 3;

QY 1 MAPITAYSQTRGLGCIITSLTGRDNQVEGEVQVIVSTATQSFATCVNGVCWTVVHGA 60

Db 1026 LAPITAYAQTRGLGCIITSLTGRDNQVEGEVQIVSTAAQTFLATCINGVCWTVVHGA 1085

QY 61 GSKTLAGKGPITOMYTNVDQLVGWQAPPAGSRLTPTCTCGSSDLVLYTRHADVIPVRRR 120

Db 1086 GTRTIASPKGPVIOMYTNVDQLVGWQAPPAGSRLTPTCTCGSSDLVLYTRHADVIPVRRR 1145

QY 121 GDSRGSLLSPRPVSVYLKSSGGPPLCPGSHAVGIFRAAVCTRGVAKAVDFVVESEMETM 180

Db 1146 GDSRGSLLSPRPISYLGKSSGGPPLCPAGHAVGIFRAAVCTRGVAKAVDFIVENLETTM 1205

QY 181 RSPVFTDNSSPAPVQTFQVAHLHAPTSGKSTKVPAAAYAAQYKVLVLPNSVAATLFGG 240

Db 1206 RSPVFTDNSSPAPVQTFQVAHLHAPTSGKSTKVPAAAYAAQYKVLVLPNSVAATLFGG 1265

QY 241 AYMSKAHGDINIRTVNRTITGAPITYSTYKFLADGGCSGGAYDIIICDECHSTDTT 300

Db 1266 AYMSKAHGDINIRTVNRTITGSPITYSTYKFLADGGCSGGAYDIIICDECHSTDTT 1325

QY 301 ILGIGTVLDOAETAGARLWVLATATPPGQSVTVPHNIEEVALSSTGEIPFFYKAIPIETI 360

Db 1326 ILGIGTVLDOAETAGARLWVLATATPPGQSVTVPHNIEEVALSSTGEIPFFYKAIPIETI 1385

QY 361 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGLDVSIVPTSGDVIWVATDALMTGFT 420

Db 1386 KGRHLIFCHSKKKKDELAALKSLGLNNAVAYYRGLDVSIVPTSGDVIWVATDALMTGFT 1445

QY 421 GDFSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRSQRRGTGRGRMGIYRFVTPG 480

Db 1446 GDFSVIDCNTCVTQTVDFSLDPTFTIETTTLPQDAVSRTQRRGRTGRKPGIYRFVAPG 1505

QY 481 ERPSGMDSSVLCBDCYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWSVFTGL 540

Db 1506 ERPSGMDSSVLCBDCYDAGCAWYELTPAETSVRLRAYLNTPLGLPVCQDHLFEWSVFTGL 1565

QY 541 THIDAHFLSOTKQAGDNPPYLVQATVCARAQAPPPSDQMWKCLIRKLPTLHGPTLL 600

Db 1566 THIDAHFLSOTKQSGENFPYLVQATVCARAQAPPPSDQMWKCLIRKLPTLHGPTLL 1625

QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVTTSTWLVGGVLAALAAAYCLTTGTVVIVGR 660

Db 1626 YRLGAVQNEVTLTHPTVKYIMTMSADLEVTTSTWLVGGVLAALAAAYCLSTGCVVIVGR 1685

QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMOLABQPKQKAIGLLOTATKQAE 720

Db 1686 IVLSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMOLABQPKQKAIGLLOTATKQAE 1745

QY 721 AAPVVEKWKRTLEAFNAKHMNFISGQYLAGLSTLPCNPAISLMAFTASITPLTTQH 780

Db 1746 IAPAVQINWQKLEVFNAKHMNFISGQYLAGLSTLPCNPAISLMAFTAAVTSPLTTSQ 1805

QY 781 TLLFNILGGWVAQOLAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYGAGVAGALVA 840

Db 1806 TLLFNILGGWVAQOLAAPGAATAFVAGLAGAAGIAGVGLKVLVDILAGYGAGVAGALVA 1865

QY 841 FKVMSGEMPTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEVAVWNRILIAFASR 900

Db 1866 FKVMSGEMPTEDLVNLLPAILSPGALVGVWCAAILRRHVGPGEVAVWNRILIAFASR 1925

QY 901 GNHVSPTHYVPESDAAARVTOILSSLTITOLLKRLHOWINEDCSTPCSGWLRDWDWIC 960

Db 1926 GNHVSPTHYVPESDAAARVTOILSSLTITOLLKRLHOWINSECTTCCSGWLRDWDWIC 1985

QY 961 TVLTDFKTLQSKLLPLPLPGVPPFFSCQYKGVWRGDMQTTCCPCQAQITGHVANGSMR 1020

Db 1986 EVLSDFKTLKAKLMPOLPGIPFVSCQYKGVWRGDMQTTCCPCQAQITGHVANGSMR 2045

QY 1021 IVGPRTCSNTHGTFPPINATTTGCTPSPAPNYSRALWRVAABEYVEVTVRGDFHYVTGM 1080

Db 2046 IVGPRTCRNMSGTFPPINATTTGCTPSPAPNYSRALWRVAABEYVEVTVRGDFHYVTGM 2105

QY 1081 TTDNVKCPQVPAPEFTEVDGVLRLHRYAPACKPLLEEVTVLGLNQYLVGSQLPCEPE 1140

Db 2106 TTDNLKCPQVPSPEFTEVDGVLRLHRYAPACKPLLEEVTVLGLNQYLVGSQLPCEPE 2165

[illegible]

661	QY	IILSGKALIPDREVLRYEEDMEECASHLPYIEQGMGLAEQFKQKAIIGLOTATKQAEA	720
1686	DB	IVLSGKPAIIPDREVLRYEEDMEECASHLPYIEQGMGLAEQFKQKALGILLQTASQAEV	1745
721	QY	AAPVVEKWRITLEAFNAKHMWNFTSGIYIAGLSTLPGNPAIASLMAFTASITSPITTOH	780
1746	DB	ITPAVQINWQKLEVFNAKHMWNFTSGIYIAGLSTLPGNPAIASLMAFTAAVTSPLTTGQ	1805
781	QY	TLLFNILGGWAAQIAPPSAASAVFGAGIAGAAGVSTGLCKVLVDIILAGYGAGVAGALVA	840
1806	DB	TLLFNILGGWAAQIAPPAAGAAATFVGAGLAGAAIGSVGLCKVLVDIILAGYGAGVAGALVA	1865

RESULT 12  
US-09-742-659-4  
; Sequence 4, Application US/09742659  
; Patent No. US20010034019A1  
; GENERAL INFORMATION:  
; APPLICANT: Hong, Zhi  
; APPLICANT: Butkiewicz, Nancy J.  
; APPLICANT: Zhong, Weidong  
; APPLICANT: Ingravallo, Paul  
; APPLICANT: Wright-Minogue, Jacqueline  
; APPLICANT: Lau, Johnson Y.  
; APPLICANT: Lemon, Stanley M.



```
QY      841  EKVMGEMSTEDLVNLLPAILLSPGALVGVVCAAILRRHVGPGEAGVOMNRLIAFASR  900
Db      1866  FKIMGEVSTEDLVNLLPAILLSPGALVGVVCAAILRRHVGPGEAGVOMNRLIAFASR  1925
QY      901  GNHVSPTHVVPESDAAARVTQLSSLTITQLLKRHLHOWINEDCSTPCSGSWLRDLWDMIC  960
Db      1926  GNHVSPTHVVPESDAAARVTQLSSLTITQLLKRHLHOWINEDCSTPCSGSWLRDLWDMIC  1985
QY      961  TVLTDFKWLQSKLLPRLPBGPFFSCQRYGKQYKVMRGDGMQITPCGAQITGHVKNRSMR  1020
Db      1986  EYLSDFKTLWKAKLMPQLPGIPFVSQCRGYRGVWRGDMGIMHTRCHCGABITGHVKNRSMR  2045
QY      1021  IVGPRTCSNTHGTTPINAYTTGCTPSPAPNYSALMRVAAEEVETVRGDFHYVTOM  1080
Db      2046  IVGPRTCSNTHGTTPINAYTTGCTPSPAPNYSALMRVAAEEVETVRGDFHYVTOM  2105
QY      1081  TTDNVKPCQVPAPEFFTEVDGVRHLRYAPACKPLLRREVTPLVGLNOYLVGSQPCPE  1140
Db      2106  TTDNLKCPQIPSPFFTELDGVRHLRFPACKPLLRREVSFVGLHYPVGSQPCPE  2165
QY      1141  PDVAVLTSMIDPSHITAEAKRRLARGSPPSLASSASQLSAPSLKATCTTNRHSDPAD  1200
Db      2166  PDVAVLTSMIDPSHITAEAKRRLARGSPPSLASSASQLSAPSLKATCTTNRHSDPAD  2225
QY      1201  LIEANLLWRQEMGGNI TRVESNKVVIIDSEPLQAEEDEREVSYPAEILRRSRKPPRAM  1260
Db      2226  LIEANLLWRQEMGGNI TRVESNKVVIIDSEPLQAEEDEREVSYPAEILRRSRKPPRAM  2285
QY      1261  PIWAPDYNPPLLESWKDPDYPVPPVHGCPLPPAKAPPPIPPRRKRTVVLSESTVSSALA  1320
Db      2286  PWARPDYNPPLVETWKKEPDYEPVHGCPLPPRSPPPPPPPPPPPPPPPPPPPPPPPPP  2345
QY      1321  ELATKTFGSSSSAVDSGTATASPPQSDG--DAGSDVESYSSMPLEGEPPDPLSDGS  1379
Db      2346  ELATKTFGSSSTSGT-TGDNNTTTSSEPAPSGCPDSDVESYSSMPLEGEPPDPLSDGS  2404
QY      1380  WSTVSEEA-SEDVVCSSSYMTGALITPCAEEFKLPINALNSLLRHHNLVATTSSR  1436
Db      2405  WSTVSSGADTEDVVCSSSYMTGALVTPCAAEOKLPINALNSLLRHHNLVATTSSR  2464
QY      1439  ASLRQKKVTFDLQVLDHRYDLVKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGY  1498
Db      2465  ACQKQKVTFDLQVLDHRYDLVKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGY  2524
QY      1499  AKOVNLSKAVNHRSVWKDLLEDETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD  1558
Db      2525  AKOVNLSKAVNHRSVWKDLLEDETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD  2584
QY      1559  LGVRVCEKMALYDVVSTLPOAVMGSSYGFQYSPGRVFEFLVNAWKAKCPMPAYDTRCF  1618
Db      2585  LGVRVCEKMALYDVVSTLPOAVMGSSYGFQYSPGRVFEFLVNAWKAKCPMPAYDTRCF  2644
QY      1619  DSTVTENDIRVEESIYQCCDLAPARQAIRSLTBLTYIGGPLTNSKNGCNGYRCRASGV  1678
Db      2645  DSTVTENDIRVEESIYQCCDLAPARQAIRSLTBLTYIGGPLTNSKNGCNGYRCRASGV  2704
QY      1679  LTTSCGNLTLCYKAAACRAAKLQDCTMLVCGDDLVIICCSAGTQDEASLRAFTAMT  1738
Db      2705  LTTSCGNLTLCYKAAACRAAKLQDCTMLVCGDDLVIICCSAGTQDEASLRAFTAMT  2764
QY      1739  RYSAPPGDPKPEYDLELITSCSSNVSAHDAKGRVYVLTEDPTPLARAAMTARHTP  1798
Db      2765  RYSAPPGDPKPEYDLELITSCSSNVSAHDAKGRVYVLTEDPTPLARAAMTARHTP  2824
QY      1799  VNSWLGNIIMVAPTILWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDPLQIQ  1858
Db      2825  VNSWLGNIIMVAPTILWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDPLQIQ  2884
QY      1859  RLHGLSFLSHSYSGEINRVASCLRLGVPLPLVRWHRARSVRARLLSQGGRAATCGKY  1918
Db      2885  RLHGLSFLSHSYSGEINRVASCLRLGVPLPLVRWHRARSVRARLLSQGGRAATCGKY  2944
```

```
QY      1919  LFNWAVRTKLTPTIPAAASQDLSSWFWAGYSGGDIYHLSLRARPRFWFWCCLLLSVGVG  1978
Db      2945  LFNWAVRTKLTPTIPAAAGRLDLSGWFTAGYSGGDIYHVSHPARFWFWCCLLLAAGVG  3004
QY      1979  IYLLPNR 1985
Db      3005  IYLLPNR 3011

RESULT 13
US-09-891-894-3
; Sequence 3, Application US/09891894
; Publication No. US20030013081A1
; GENERAL INFORMATION:
; APPLICANT: Olson, William
; APPLICANT: Maddon, Paul
; TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C VIRUS INFI
; FILE REFERENCE: 2048/64896/JPW/SHS
; CURRENT APPLICATION NUMBER: US/09/891,894
; CURRENT FILING DATE: 2001-06-26
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 3011
; TYPE: PRK
; ORGANISM: hepatitis c virus
US-09-891-894-3
```

```
Query Match      90.1%; Score 9430.5; DB 10; Length 3011;
Best Local Similarity 87.8%; Pred. No. 0;
Matches 1745; Conservative 136; Mismatches 103; Indels 3; Gaps 3;

QY      1  MAPITAYSQQRGLLIGCIITSLTCGRDRNOVEGEVOWVSTATQSLATCVNGVCWTVYHGA  60
Db      1026  LAPITAYAAQTRGLLIGCIITSLTGEDKNQVEGEVQIVSTATQTFLATCINGVCWTVYHGA  1085
QY      61  GSKTLAGPKBITQMYINVDQDLVGMQAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRR  120
Db      1086  GTRTIASPKGPIQMYINVDQDLVGMQAPPGARSLTPCTCGSSDLYLVTRHADVIPVRRR  1145
QY      121  GDSRGSLLSPRPVSVLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM  180
Db      1146  GDSRGSLLSPRPISYLLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVENLETTM  1205
QY      181  RSPVFTDNSSPPAVPQFQVAHLHAPTGSKGSKTVPAAYAAQGYKVLVLPNSVAATLGG  240
Db      1206  RSPVFTDNSSPPAVPQFQVAHLHAPTGSKGSKTVPAAYAAQGYKVLVLPNSVAATLGG  1265
QY      241  AYMSKAHGDINIRTGVRTITTGAPITVSTYKFLADGGCSGGAYDIIICDECHSTDST  300
Db      1266  AYMSKAHGVDPNIRTGVRTITTGSPITVSTYKFLADGGCSGGAYDIIICDECHSTDATS  1325
QY      301  ILGIGTVLDOAETAGARLVLATATPGSVTVPHNIEEVALSSTGEIPFGKAIPIETI  360
Db      1326  ILGIGTVLDOAETAGARLVLATATPGSVTVPHNIEEVALSSTGEIPFGKAIPIEVI  1385
QY      361  KGRHLLIFCHSKKKCDELAALKSLGLNAVAYRGLDVSVIPTSGDVIWVATDALMTGFT  420
Db      1386  KGRHLLIFCHSKKKCDELAALKSLGLNAVAYRGLDVSVIPTSGDVIWVATDALMTGFT  1445
QY      421  GDFDSVIDCNTCTVQTVDFSLDPTFTTETTTVPQDAVSRQRRGRTGRGMGIYRFVTPG  480
Db      1446  GDFDSVIDCNTCTVQTVDFSLDPTFTTETTTVPQDAVSRQRRGRTGRGKPGIYRFVAPG  1505
QY      481  ERSGMFDSSLCECYDAGCAWYELTAPETSRLRAVLNTPGLPVCODHLEFESVETGL  540
Db      1506  ERSGMFDSSLCECYDAGCAWYELTAPETSRLRAVLNTPGLPVCODHLEFEGVFTGL  1565
QY      541  THIDAHFLSQTQKAGDNFPYLVAYQATVCARAQAPPPSWDMKCLRLKPTLHGPTPL  600
Db      1566  THIDAHFLSQTQKSGENFPYLVAYQATVCARAQAPPPSWDMKCLRLKPTLHGPTPL  1625
QY      601  YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLITGTSVVIVGR  660
```

Db 1626 YRLGAVQNEVTLTHPTIKITMTCMSADLEWVTSWVLVGGVLAALAAAYCLSTCCVIVGR 1685  
 QY 661 IILSGPAIIPDREVLVREFDEMEECASHLPYIEQGMQALAEQKQKAIGLQATKQAEA 720  
 Db 1686 IIVLSGPAIIPDREVLVQFDEMEECASHLPYIEQGMWLAELQKQKALGILLQATSRQAEV 1745  
 QY 721 AAPVSEKRTLEAFWAKHWNFIISGLOYLAGLSTLPGNPAIASLMAFTASISPLTTOH 780  
 Db 1746 ITPAVQTNWQKLEVFVAKHWNFIISGLOYLAGLSTLPGNPAIASLMAFTAAVTSPLTTOG 1805  
 QY 781 TILFNILGGHVAQAALAPPASAFVAGIAGAAGVSGIGLKVLDIILAGYAGAGVAGALVA 840  
 Db 1806 TILFNILGGHVAQAALAPGATAFVAGLAGAAGVSGIGLKVLDIILAGYAGAGVAGALVA 1865  
 QY 841 FYKMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGGEGAGVQNMNLLIAPASR 900  
 Db 1866 FKIMSGEVPTEDLVNLLPAILSPGALVGVCAAILRRHVGGEGAGVQNMNLLIAPASR 1925  
 QY 901 GNEVSPHYVPESDAARVTCILSSLTITQLKRLHOWINEDCSTPCSGSWLDRVMDWIC 960  
 Db 1926 GNEVSPHYVPESDAARVTCILSSLTITQLKRLHOWINEDCSTPCSGSWLDRVMDWIC 1985  
 QY 961 TVLTDFTKTLQSKLPLPGVPFPFCQYKGYVMRGDGMQTTCPGCAQITGHVYKSGMR 1020  
 Db 1986 EVLSDFKTLKALXKMLPQLPGIPFVSCQYRGVYVMRGDGMHTRCHCGAEITGHVYKGTMR 2045  
 QY 1021 IVERTCSTNHWGTFPPINAVTGTGCTSPAPNYSRALWRVAABEYVEVTRVGFHYHVTGM 1080  
 Db 2046 IVPRTCRNWSGTFPPINAVTGTGCTPLPAPNKFALWRVSAEYVEIRRVGFHYHVTGM 2105  
 QY 1081 TTNVXKCPQVPAPEPTEVDGLHRYAPACKPLLEETVGLNOLYVLSGOLPCEPE 1140  
 Db 2106 TTNMLKCPQIPSEPFTEVDGLHRYAPACKPLLEETVGLNOLYVLSGOLPCEPE 2165  
 QY 1141 PDVAVLTSMLTDSHITAEAKRLARGSPSPSLASSASQSLAPSLKATCTTTRHDSPAD 1200  
 Db 2166 PDVAVLTSMLTDSHITAEAGRELARGSPSPMASSASQSLAPSLKATCTANHDSPAD 2225  
 QY 1201 LIBANLWRQEMGNITRVSEKNVILDSPEIQAEDEREVSVPAAEILRRSRKFFRAM 1260  
 Db 2226 LIBANLWRQEMGNITRVSEKNVILDSFDPVLAEDEREVSVPAAEILRRSRKFFRAM 2285  
 QY 1261 PIWAPDYNPLLESWKDPYVPPVHGCPLPPAKAPPPIPPRRKRTVILSESIVSSALA 1320  
 Db 2286 PWARPDYNPLVETWKKDPYVPPVHGCPLPPRSPPPVPPRKRIVLVLESTLSTALA 2345  
 QY 1321 ELATKTFGSSSAVDSGTASPDQSDDG - DAGSDVESYSSMPLEGEFGDPDLSGDS 1379  
 Db 2346 ELATKTFGSSSTSGI - TGDNTTTSSEAPSGCPCPDSDVESYSSMPLEGEFGDPDLSGDS 2404  
 QY 1380 WSTVSEBA - SEDVVCCSMSTWTGALLTPCAAEETKLPINALSNLLRHNLVYATTSRS 1438  
 Db 2405 WSTVSSGADTEDVVCCSMSTWTGALVTPCAAEQKLPINALSNLLRHNLVYATTSRS 2464  
 QY 1439 ASLRQKVTDFRLQVLDHVDVLEKMAKASTVKAKLLSVEEACKLTPPHSARSKTGYG 1498  
 Db 2465 ACQKQKVTDFRLQVLDHVDVLEKMAKASTVKAKLLSVEEACKLTPPHSARSKTGYG 2524  
 QY 1499 AKQVRNLSKAVNHRSVWKLLEDTETPIDTTIMAKNEVCVQPEKGRKPARLIVFPD 1558  
 Db 2525 AKQVRCHARKAVAHINSVWKLLEDSVTPIDTTIMAKNEVCVQPEKGRKPARLIVFPD 2584  
 QY 1559 LGVRVCEKMAIYDVVSTLPQAVMSSYGFQYSPQRVEFLVNAWKAKCPMGFADTRCF 1618  
 Db 2585 LGVRVCEKMAIYDVVSTLPQAVMSSYGFQYSPQRVEFLVNAWKAKCPMGFADTRCF 2644  
 QY 1619 DSTVTENDIRVEESYOCCLDAPARQAIRSLTERLYIGGLPTNSKONCGYRRCRASGV 1678  
 Db 2645 DSTVTESDIRVEEAIYQCCLDAPARVAIKSLTERLYVGGPLTNSRGENCGYRRCRASGV 2704  
 QY 1679 LTTSCGNTLTCTYKAAAAACRAAKIQDCTMLVCGDDLVVICSSAGTQDEASIRAPTEAMT 1738

Db 2705 LTTSCGNTLTCTYKAAAAACRAAKIQDCTMLVCGDDLVVICSSAGTQDEASIRAPTEAMT 2764  
 QY 1739 RYSAPPDGPDKBEYDLELITSCSSNVSAHDASGKRYVYLTROPTTTLARAAWETAHTP 1798  
 Db 2765 RYSAPPDGPDPQBYDLELITSCSSNVSAHDAGKRYVYLTROPTTTLARAAWETAHTP 2824  
 QY 1799 VNSWLGNIIIMYAPTILWARMILMTHFFSIIAQAQLEKALDCQIYGACYSIEPLDLPIIQ 1858  
 Db 2825 VNSWLGNIIIMYAPTILWARMILMTHFFSVLIARQLEALNCIYGACYSIEPLDLPIIQ 2884  
 QY 1859 RLHGLSAFSLHSYSPGEINRVASCLRKLGVPPPLRVWRHARSVRARLLSOGGAAATCGKY 1918  
 Db 2885 RLHGLSAFSLHSYSPGEINRVAAACLRKLGVPPPLRVWRHARSVRARLLSOGGAAATCGKY 2944  
 QY 1919 LFNWAVRTKLTPIPAAASOLDLSSWFMVAGYSGDIIYHLSRARPRFWMFMCLLLLSVGVG 1978  
 Db 2945 LFNWAVRTKLTPIPAAAGRLDLSGWNFTAGYSGDIIYHLSVSHARPRFWMFMCLLLLSVGVG 3004  
 QY 1979 IYLLPNR 1985  
 Db 3005 IYLLPNR 3011

RESULT 14  
 US-10-184-150-3  
 ; Sequence 3, Application US/10184150  
 ; Publication No. US20030134297A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Olson, William  
 ; APPLICANT: Maddon, Paul  
 ; TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C VIRUS IN  
 ; FILE REFERENCE: 2048/64896-A/JPW/MAF/DJK  
 ; CURRENT APPLICATION NUMBER: US/10/184,150  
 ; CURRENT FILING DATE: 2002-12-10  
 ; PRIOR APPLICATION NUMBER: 09/891,894  
 ; PRIOR FILING DATE: 2001-06-26  
 ; NUMBER OF SEQ ID NOS: 3  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 3  
 ; LENGTH: 3011  
 ; TYPE: PRT  
 ; ORGANISM: Hepatitis C virus  
 US-10-184-150-3

Query Match 90.1%; Score 9430.5; DB 14; Length 3011;  
 Best Local Similarity 87.8%; Pred. No. 0;  
 Matches 1745; Conservative 136; Mismatches 103; Indels 3; Gaps 3;

QY 1 MAPITAYSQQTRGLLGCITSLTGRDRNQVEGEVQVYVSTATQSPATCNGVCMVTVHGA 60  
 Db 1026 LAPITAYAQTRGLLGCITSLTGRDRNQVEGEVQVYVSTATQSPATCNGVCMVTVHGA 1085  
 QY 61 GSKTLAPGKPIQMYTNVDQDLVGVQWQAPPGARSLLTCTCGSSDLVLTSHADVIPVRR 120  
 Db 1086 GTRTIASPKGPVQMYTNVDQDLVGVQWQAPPGARSLLTCTCGSSDLVLTSHADVIPVRR 1145  
 QY 121 GDSRGSLLSPRVSYLKSGSGGGLLPCPSGHAUVGFRAAVCTRGVAKAVDFVFPVBSMETTM 180  
 Db 1146 GDSRGSLLSPRVSYLKSGSGGGLLPCPSGHAUVGFRAAVCTRGVAKAVDFVFPVBSMETTM 1205  
 QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAAQYKVLVLPNSVAATLGF 240  
 Db 1206 RSPVFTDNSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAAQYKVLVLPNSVAATLGF 1265  
 QY 241 AYMSKAGIDPNIRTGVRTITTCGAPITYSTYKFLADGGCGGGAYDIIICDECHSTDTT 300  
 Db 1266 AYMSKAGIDPNIRTGVRTITTCGAPITYSTYKFLADGGCGGGAYDIIICDECHSTDTT 1325  
 QY 301 ILGIGTVLDQAEATAGARLVVLTATPPGSVTVPHNIEEVALSSTGEIPFYKAIPIETI 360  
 Db 1326 ILGIGTVLDQAEATAGARLVVLTATPPGSVTVPHNIEEVALSSTGEIPFYKAIPIETI 1385  
 QY 361 KGRHLLIFCHSKKKKDELAALKSLGLNNAVAYRGLDVSIVPTSGDVIIVATDALMTGFT 420

Db	1386	KGGSHLLFCHSKKKKDELAALKVALGINAVAYRGLDVSIVPTSGDVVVVSTDALMTGFT	1444
Qy	421	GDFDSVIDCNTCVTQTVDSDLDPFTTIETTTVPDAVSRSORRGRTGRGMGIYRFVTPG	480
Db	1446	GDFDSVIDCNTCVTQTVDSDLDPFTTIETTTLPDAVSRTORRORTGRGPKIYRFVAPG	1505
Qy	481	ERPSGMFDSSVLCECYDAGCAWYBELTPAETSVRURAYLINTPGLPVCDHLEFWESVFTEL	540
Db	1506	ERPSGMFDSSVLCECYDAGCAWYBELTPAETTIVRLRAYMNTPGLPVCDHLEFWESVFTEL	1565
Qy	541	THIDAHFLSOTKOAGDNFFYLVAOATVCARAOAAPPSPWDOMWKILRLKPTLHGPPPLL	600
Db	1566	THIDAHFLSOTKOSGENFPYLVAOATVCARAOAAPPSPWDOMWKILRLKPTLHGPPPLL	1625
Qy	601	YRLGAVQNEVTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVIUGR	660
Db	1626	YRLGAVQNEVTLHPITKIYIMTMSADLEVVTSTWLVGGVLAALAAAYCLTGCCVVIUGR	1685
Qy	661	IILSGKPAILPDREVLVREFDEMEECASHLPYIEQGMLQAEPFKQKAIGLLQATKQAEA	720
Db	1686	IIVLSGKPAILPDREVLVQEDEFEMEESCQHLPYIEQMMLAEQFKQKALGILLQATSRQAEV	1745
Qy	721	AAPVESKWTLEAFWAKHMWNFIISGYLAGLSTLPGNPAIASLMFTAASITSPLTTQH	780
Db	1746	ITPAVQTNWOKLEVFWAKHMWNFISGYLAGLSTLPGNPAIASLMFTAATAVTSPLTTGQ	1805
Qy	781	TLLFNILGGWVAQLAPPSAASFVGAGIAGAAGVSGIGLKVLVDIILAGYGAGVAGALVA	840
Db	1806	TLLFNILGGWVAQLAAPAGTAATFVGAGLAGAAGSVGLGKVLDIILAGYGAGVAGALVA	1865
Qy	841	EKMVGEMSPTEDLVNLLPAILPGALVGVWCVAAILRRHVGPGEAVOMNRLLIAFASR	900
Db	1866	EKIMSGEVPTEDLVNLLPAILSPGALVGVWCVAAILRRHVGPGEAVQWMNRLLIAFASR	1925
Qy	901	GNHVSPTHYPESDAAARVTCIISSLTITOLLKRIHQINEDCSTPCSGSWLRDWDWIC	960
Db	1926	GNHVSPTHYPESDAAARVTAIISSLTVTOLLKRLHQIWISSBECTTCPGSGSWLRDIWDWIC	1985
Qy	961	TVLTDFKTWLQSLLPRLPGVPPFSCQRGKGVWRGDGIMOTTCPCGAQITGHVKNGSMR	1020
Db	1986	EVLSDFKTWLXKLMPOLPGIPFVSCQRGVGWRGDGMHTKCHCAEITGHVKNGTWR	2045
Qy	1021	IVGPRTCSTNHGTFPINAYTTGCTPSPAPNYSRALRWAAEEYVEVTRVGFHYVTGM	1080
Db	2046	IVGPRTCSTNHGTFPINAYTTGCTPLPAPNKFALRWAAEEYVEIRRVGDFHYVSGM	2105
Qy	1081	TTDNVCKCOVPAPPEFTEVDGVLRLHRYAPACKPLREEVTVLVLNQYLVGSQLPCBPBE	1140
Db	2106	TTNLNCKCQIPSEPFTELDGVLRLHRFAFPCKPELLREEVSFRVGLHEYPVGSQLECEPE	2165
Qy	1141	PDVAVLTSLMLTDPSHITAETAKRRLARGSPPSLASSASOLSAPSLKATCTTRHDSPPDAD	1200
Db	2166	PDVAVLTSLMLTDPSHITAEAAAGRLARGSPPSWASSASOLSAPSLKACTANHSPPDAE	2225
Qy	1201	LIEANLLWROEMCGNITRVESENKVVILOSFEPLQAEDEEREVSVPAAELRLRSRKFPAM	1260
Db	2226	LIEANLLWROEMCGNITRVESENKVVILOSFDPLVAEDEEREVSVPAAELRLRSRRFARAL	2285
Qy	1261	PIWARPDYNPPLLESKWDPDYPPVNVHGCPLPAKAPPPIPPRRKXTVILSSTVSSALA	1320
Db	2286	PVARPDYNPPLVETWKKPDYEPVNVHGCPLPPRPSPFPVPPPPKKTVMILTTESTLSTA	2345
Qy	1321	ELATKTFGSSSESAVDSGTATASDPQSDDG-DAGSDVESYSMPPLEGEGDPLSDGS	1379
Db	2346	ELATKFGSSSTSGI-TGDNTTTTSEPAAGCGCPDSDVESYSMPPLEGEGDPLSDGS	2404
Qy	1380	WSTVVSSEA-SEDVVCCSMSYTWTGALITPCAABETKLINALSNSLLRHNLVYATTSSR	1438
Db	2405	WSTVSSGAUTEDVVCCSMSYMTGALVTPCAAEEQKLINALSNSLLRHNLVYSTTSSR	2464
Qy	1439	ASLRQKKVTFDRLQVLDHVRDVLKEMKAKASTVKALLSVEBAACKLTPPHSARSKEGYG	1498

Db	2465	ACQKQKVTFDRQLVDLSHYQDVLKRVKAAASKYKANLLSVEEACSLTPPHSAKSKEGYG	2524
Qy	1499	AKDVRLSSKAVNHRSWVKDLLEDTEPTDITTIMAKNEVFCVQPEKGGKRPKRLIIVFPD	1558
Db	2525	AKDVRCHARKAVAHNSWVKDLLEDSTVPTDITTIMAKNEVFCVQPEKGGKRPKRLIIVFPD	2584
Qy	1559	LGVRCCKMALYDVVSTLPOAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFAYDTRCF	1618
Db	2585	LGVRCCKMALYDVVSKLPLAVMGSSYGFQYSPGQRVEFLVQAWKSKTPEMGFSYDTRCF	2644
Qy	1619	DSTVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQNGCYRRCRASGV	1678
Db	2645	DSTVTESDIRTEEAIIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGNGCYRRCRASGV	2704
Qy	1679	LTTCGNTLTCYLKAAACAAKLODCTMLVCGDDLWICESAGTOQDEASLRAFTAMT	1738
Db	2705	LTTCGNTLTCYIKARACRAAGLODCTMLVCGDDLWICESAGVGDAAASLRAFTAMT	2764
Qy	1739	RYSAPPGDPPKPEYDLELITSCSNVSVAHDSAGKRVYYLTRDPTTFLAPAAWETARHTP	1798
Db	2765	RYSAPPGDPPQPEYDLELITSCSNVSVAHGDAGKRVYYLTRDPTTFLAPAAWETARHTP	2824
Qy	1799	VNSWLGNIIMVAPTLWARMILMTHFFSILLAQOLEKALDCQYIGACYSTIEPLDLPPIIQ	1858
Db	2825	VNSWLGNIIMFAPTLWARMILMTHFFSVLIARDQLEQALNCEIYGACYSTIEPLDLPPIIQ	2884
Qy	1859	RLHGLSAFSLHSYSPGSEINRVASCLRKLGVPPLRVWHRHARSVRARLLSOGGAAATCGKY	1918
Db	2885	RLHGLSAFSLHSYSPGSEINRVAACLRKLGVPPPLURAWHRHARSVRARLLSOGGAAATCGKY	2944
Qy	1919	LFNWAVRTKLKLTPIPAASOLDLSWVAGYSGGDIYHSLSRAPRWFVWMLLLLSVGVG	1978
Db	2945	LFNWAVRTKLKLTPIAAAGRLDLSGWFTAGYSGGDIYHVSHPRPWFVFCLLLLAAGVG	3004
Qy	1979	IYLLPNR 1985	
Db	3005	IYLLPNR 3011	
RESULT 15			
US-10-328-997-3			
; Sequence 3, Application US/10328997			
; Publication No. US20030232745A1			
; GENERAL INFORMATION:			
; APPLICANT: Olson, William			
; APPLICANT: Madden, Paul			
; TITLE OF INVENTION: USES OF DC-SIGN AND DC-SIGNR FOR INHIBITING HEPATITIS C V			
; FILE REFERENCE: 2048/64896-B/JPW/NAF/DJK			
; CURRENT APPLICATION NUMBER: US/10/328,997			
; CURRENT FILING DATE: 2002-12-24			
; PRIOR APPLICATION NUMBER: 09/891,894			
; PRIOR FILING DATE: 2001-06-26			
; PRIOR APPLICATION NUMBER: 10/184,150			
; PRIOR FILING DATE: 2002-06-26			
; NUMBER OF SEQ ID NOS: 3			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 3			
; LENGTH: 3011			
; TYPE: PR			
; ORGANISM: Hepatitis C virus			
US-10-328-997-3			

```

Query Match      90.1%; Score 9430.5; DB 14; Length 3011;
Best Local Similarity 87.8%; Pred. No. 0;
Matches 1745; Conservative 136; Mismatches 103; Indels 3; Gaps 3;

Qy 1 MAPITAYSOOTRGLGCIITSLTGDRNQVEGEVQVSTQSFATCNGVCWTIVYHGA 60
   :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1026 LAPITAYAQOTRGLGCIITSLTGDRKQVEGEVQIVSTATQTFLATCINGVCWTIVYHGA 1085

Qy 61 GSKTLAPGKPIOMYNTVDQLVQWAPPGARSLTPCTCGSSDLYLVTRIADVIVRRR 120
   :|::|:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 1086 GRTITASPKGPVIOMYNTVDQLVQWAPPQGSRLTPCTCGSSDLYLVTRIADVIVRRR 1145

```

QY	121	GDGRGILLSPRPVSYLYKSGSGGELLCPGSHAVGI FRAAVCTRGVAKAVDPVPVSMETTM	180
Db	1146	GDGRGILLSPRPVSYLYKSGSGGELLCPGSHAVGLFRAAVCTRGVAKAVDPVFPVENLETTM	1205
QY	181	RSVPFTDNGSPPAVPOTFQVAHLHAPTGSKSKTKVPAAYAAQGVKVLVLPNSVAAATLGF	240
Db	1206	RSVPFTDNGSPPAVPOTFQVAHLHAPTGSKSKTKVPAAYAAQGVKVLVLPNSVAAATLGF	1265
QY	241	AYMSKAHGIDPNIRTVGRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT	300
Db	1266	AYMSKAHGIDPNIRTVGRTITTCAPITYSTYKFLADGGCGGAYDIIICDECHSDATS	1325
QY	301	ILGIGTVLDOAETAGARLVVLAATATPPGTVTPHPNIEVALSTGEIPIFYKAIPIETI	360
Db	1326	ILGIGTVLDOAETAGARLVVLAATATPPGTVTPHPNIEVALSTGEIPIFYKAIPIEVI	1385
QY	361	KGGRHLIFCHSKKCDLAAKLSGLNNAVAYVRGLDVSVIPTSGDVIVVATDALMTGFT	420
Db	1386	KGGRHLIFCHSKKCDLAAKLSGLNNAVAYVRGLDVSVIPTSGDVIVVATDALMTGFT	1445
QY	421	GDPSVIDNCNTVQTQVDFSLDPTFTIETTTPQDAVRSQRRTGRGRMGIYRFVTPG	480
Db	1446	GDPSVIDNCNTVQTQVDFSLDPTFTIETTTPQDAVRSQRRTGRGRMGIYRFVAPG	1505
QY	481	ERPSGMPDSSVLCEDYDAGCAWELTPAETSURLRAYLNTPLPVCODHLEFWESVETGL	540
Db	1506	ERPSGMPDSSVLCEDYDAGCAWELTPAETVRLRAYMNTPLPVCODHLEFWEGVETGL	1565
QY	541	THIDAFHLSQTKAGDNFPVLYAVQATVCARAQAPPSQWQMKLLRLKPTLHGPTPLL	600
Db	1566	THIDAFHLSQTKAGDNFPVLYAVQATVCARAQAPPSQWQMKLLRLKPTLHGPTPLL	1625
QY	601	YRLGAVQNEVTTHTPIKYIMACMSADLEVVTSTWLVGVGLAALAAAYCLTTSQSVIYGR	660
Db	1626	YRLGAVQNEVTTHTPIKYIMTMSADLEVVTSTWLVGVGLAALAAAYCLTSGCWIVGR	1685
QY	661	IILSGKPAIIPDRHVLRFDEMEECASHLPYIQOQGLAEQFKQKAIQLIQTAKQEA	720
Db	1686	IIVLSGKPAIIPDRHVLRFDEMEECASHLPYIQOQGLAEQFKQKAIQLIQTASQAEV	1745
QY	721	AAPVVEKSWTLEAFWAKHWNFTSGIYLAGLSTLPGNPAIASLMAFTASITSLPTQH	780
Db	1746	ITPAVQTNQKLEFVWAKHWNFTSGIYLAGLSTLPGNPAIASLMAFTAAVTSPLTQ	1805
QY	781	TLLENILGGWAAQAPPSAASAFVAGIAGAAVSGI GLGVLDIILAGYGAGVAGALVA	840
Db	1806	TLLENILGGWAAQAPPSAASAFVAGIAGAAVSGI GLGVLDIILAGYGAGVAGALVA	1865
QY	841	FKVMSGEMPTEDANILLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNNRLIAFASR	900
Db	1866	FKVMSGEMPTEDANILLPAILSPGALVGVVCAAILRRHVGPGEVAVQWNNRLIAFASR	1925
QY	901	GNHVSPTHYPESDAAARVTOILSSLTITOLLKELHOWINEDCTPCSGSWLRDWDWIC	960
Db	1926	GNHVSPTHYPESDAAARVTOILSSLTITOLLKELHOWISECTTPCSGSWLRDWDWIC	1985
QY	961	TVLDFKTLQSKLLPRLPGVPFPCQGRYKGVWRGDGIMOTTCPGCAQITGHVKNQSMR	1020
Db	1986	EVLSDFKTLQSKLLPRLPGVPFPCQGRYKGVWRGDGIMOTRCHCAGAEITGHVKNQSMR	2045
QY	1021	IVGPRCTSNWHTGTFPINAYTTGCTPSPAPNSRALWRVAEEVEVTRVGDFFHYVTGM	1080
Db	2046	IVGPRCTSNWHTGTFPINAYTTGCTPPLPAPNYFALWRVAEEVEVTRVGDFFHYVSGM	2105
QY	1081	TTDNVVCPCQVPAPEFFTEVDGVRHLRYAPACKELLREEVTFVLGNQYLVGSQLPCEPE	1140
Db	2106	TTDNVVCPCQVPAPEFFTEVDGVRHLRYAPACKELLREEVTFVLGNQYLVGSQLPCEPE	2165
QY	1141	PDVAVLTSMLTDPSSHITATKRLRARGSPPSLASSASQLSAPSLKATCTTRHDSPPAD	1200
Db	2166	PDVAVLTSMLTDPSSHITATKRLRARGSPPSLASSASQLSAPSLKATCTANHDSPPAE	2225

Search completed: December 8, 2004, 12:36:07  
Job time : 138 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:19:47 ; Search time 61 Seconds  
(without alignments)  
3130.988 Million cell updates/sec

Title: US-09-576-989-3

Perfect score: 10465

Sequence: 1 MAPITAYSQQRGLLGCIIT.....FMWCLLLLSVGVIGIYLLPNR 1985

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_79: \*  
1: PIR1: \*  
2: PIR2: \*  
3: PIR3: \*  
4: PIR4: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	10170	97.2	3010	1 GNVVTC	genome polyprotein
2	10159	97.1	3010	1 A45573	genome polyprotein
3	10109	96.6	3010	1 GNVVJC	genome polyprotein
4	10069	96.2	3010	1 S18030	genome polyprotein
5	10066	96.2	3010	1 GNVVTV	genome polyprotein
6	9417.5	90.0	3011	1 GNVVVC	genome polyprotein
7	9398.5	89.8	3011	1 S40770	genome polyprotein
8	9266.5	88.5	3011	1 GNVVCH	genome polyprotein
9	8412.5	80.4	3014	1 JC5620	genome polyprotein
10	8037.5	76.8	3033	1 JQ1303	genome polyprotein
11	7959.5	76.1	3033	1 GNVVJ8	genome polyprotein
12	2383	22.8	492	2 PS0326	genome polyprotein - hepa
13	2098	20.0	876	2 PC2219	polyprotein - hepa
14	2078.5	19.9	3005	2 T08841	polyprotein - hepa
15	1980.5	18.9	2970	2 T08839	polyprotein - marm
16	1921.5	18.4	386	2 S68016	ATPase/RNA helicas
17	1855	17.7	365	2 J00879	NS5 protein - hepa
18	1755	16.8	365	2 J00880	NS5 protein - hepa
19	1555	14.9	716	2 JQ1366	polyprotein - hepa
20	1468	14.0	874	2 J00883	genome polyprotein
21	1457	13.9	874	2 J00881	genome polyprotein
22	1112	10.6	200	2 PQ0246	polyprotein (clone
23	1097	10.5	200	2 PQ0245	polyprotein (clone
24	1091	10.4	216	2 S21337	genome polyprotein
25	1048.5	10.0	1435	2 T01075	polyprotein - hepa
26	1036	9.9	194	2 A54317	probable nonstruct
27	1034	9.9	194	2 S06067	nonstructural prot
28	986	9.4	200	2 PQ0248	polyprotein (clone
29	983	9.4	184	2 A61196	genome polyprotein

#### ALIGNMENTS

##### RESULT 1

GNVWTC

genome polyprotein - hepatitis C virus

N:Contains: capsid protein C; envelope protein M; hepatitis M; hepatitis (EC 3.4.21.98) (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 09-Jul-2004

C:Accession: A38465

R:Takamizawa, A.; Mori, C.; Fuke, I.; Manabe, S.; Murakami, S.; Fujita, J.; Onishi, E.; J

J. Virol. 65, 1105-1113, 1991

A:Title: Structure and organization of the hepatitis C virus genome isolated from human c

A:Reference number: A38465; MUID:91140698; PMID:1847440

A:Accession: A38465

A:Molecule type: Genomic RNA

A:Residues: 1-3010 <TAK>

A:Cross-references: UNIPROT:P26663; EMBL:M58335; NID:g329770; PID:AAA72945.1; PID:g32977

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructural

F:2-115/Product: capsid protein C #status predicted <CPC>

F:116-191/Product: envelope protein M #status predicted <BPM>

F:192-389/Product: major envelope protein E #status predicted <ME>

F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

F:1007-1615/Product: hepatitis C virus genome polyprotein

F:1312-1317/Region: nucleotide-binding motif A (P-loop)

F:1316-1319/Region: DEHX motif

F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>

F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>

F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

F:196,209,234,250,305,325,417,423,430,448,532,540,556,576,623,645,1213,1255,2041,2077,224

##### Query Match

Best Local Similarity 97.2%; Score 10170; DB 1; Length 3010;

Matches 1919; Conservative 36; Mismatches 30; Indels 0; Gaps 0;

Qy 1 MAPITAYSQQRGLLGCIITSLTGRDNQVEGVQVWSTATQSFATCVCNGVCTVYHGA 60

Db 1026 LAPITAYSQQRGLLGCIITSLTGRDNQVEGVQVWSTATQSFATCVCNGVCTVYHGA 1085

Qy 61 GSKTLAGKPGITQMTYTNVDQDLVGVQAPPGARSLTPTCTGSSDLVLTTRHADVIPVRR 120

Db 1086 GSKTLAAPKPGITQMTYTNVDQDLVGVQAPPGARSLTPTCTGSSDLVLTTRHADVIPVRR 1145

Qy 121 GDSRGSLLSRPVSYLEKSSGGPLLCPSGHVGFRAVCTRGVAKAVDFPVESMETTM 180

Db 1146 GDSRGSLLSRPVSYLEKSSGGPLLCPSGHVGFRAVCTRGVAKAVDFPVESMETTM 1205

Qy 181 RSPVFTDSSPPAVPQTFQVAHLHAPTGSKGSKTVPAAVAAQGYKVLVLPNSVAATLGF 240

Db 1206 RSPVFTDSSPPAVPQTFQVAHLHAPTGSKGSKTVPAAVAAQGYKVLVLPNSVAATLGF 1265

QY 241 AYMSKAGHDIPNIRTCVRIITTCAPITYSTYKFLADGGCGSGAYDIIICDCBSTDST 300  
Db 1266 AYMSKAGHDIPNIRTCVRIITTCAPITYSTYKFLADGGCGSGAYDIIICDCBSTDST 1325  
QY 301 ILGIGTVLDOAETAGARLVVLAATATPGSVTVVPHNIEEVALSGTGEIPYKAIPIETI 360  
Db 1326 ILGIGTVLDOAETAGARLVVLAATATPGSVTVVPHNIEEVALSGTGEIPYKAIPIEA 1385  
QY 361 KGRHLIFCHSKKKCDKDELAALKSLGLNNAVAYVYRGDLVSVIPTSGDVVVVATDALMTGFT 420  
Db 1386 KGRHLIFCHSKKKCDKDELAALKSLGLNNAVAYVYRGDLVSVIPTSGDVVVVATDALMTGFT 1445  
QY 421 GDFSDVLDNCTCYTQTVDFSLDFTETTTTVPQDAVRSORGRGGMGIYRFVTPG 480  
Db 1446 GDFSDVLDNCTCYTQTVDFSLDFTETTTTVPQDAVRSORGRGGMGIYRFVTPG 1505  
QY 481 ERPSGMFDSVLCBCEYDAGCAWYELTPAETSRLRAYLNTPLPVCQDHLFEWESVFTGL 540  
Db 1506 ERPSGMFDSVLCBCEYDAGCAWYELTPAETSRLRAYLNTPLPVCQDHLFEWESVFTGL 1565  
QY 541 THIDAHFLSOTKQAGNFPVLVYQATVCARAQAPPSWDQMKCLIRLKPRTLHGTPPLL 600  
Db 1566 THIDAHFLSOTKQAGNFPVLVYQATVCARAQAPPSWDQMKCLIRLKPRTLHGTPPLL 1625  
QY 601 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSAVTVGR 660  
Db 1626 YRLGAVQNEVTTTHPIITKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSAVTVGR 1685  
QY 661 IILSGKPAITPDREVLRYEEDMEBCASHLPYIEOGMQLAEQFKOKAIGLLQATKQAEA 720  
Db 1686 IILSGKPAITPDREVLRYEEDMEBCASHLPYIEOGMQLAEQFKOKAIGLLQATKQAEA 1745  
QY 721 AAPVWESKWTLEAFWAKHWNFTISGIXYLAGLSTLPGNPAIASLMFTAITSPLTTQH 780  
Db 1746 AAPVWESKWTLEAFWAKHWNFTISGIXYLAGLSTLPGNPAIASLMFTAITSPLTTQH 1805  
QY 781 TLLFNILGGWVAQALAPSAASAVGAGIAGAAVSTGLGKVLVDIILAGYAGVAGALVA 840  
Db 1806 TLLFNILGGWVAQALAPSAASAVGAGIAGAAVSTGLGKVLVDIILAGYAGVAGALVA 1865  
QY 841 FKWSGEMPSDEDLVNLPAISLPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAPASR 900  
Db 1866 FKWSGEMPSDEDLVNLPAISLPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAPASR 1925  
QY 901 GNHVSPTHYVPESDAARVTOILSSLIITQLKRLHOWINEDECSTPCSGSLRVDWDMIC 960  
Db 1926 GNHVSPTHYVPESDAARVTOILSSLIITQLKRLHOWINEDECSTPCSGSLRVDWDMIC 1985  
QY 961 TVLTDFKTLQSKLLPRLPGVPFSCORGKYGWVRGDIQMTTCPCGAQITGHVKNQSMR 1020  
Db 1986 TVLTDFKTLQSKLLPRLPGVPFSCORGKYGWVRGDIQMTTCPCGAQITGHVKNQSMR 2045  
QY 1021 IVPRTCSNTHWGHFPINAYTTGCTSPAPNYSRALRWAAEYVEVTRVGDFFHYVTGM 1080  
Db 2046 IVPRTCSNTHWGHFPINAYTTGCTSPAPNYSRALRWAAEYVEVTRVGDFFHYVTGM 2105  
QY 1081 TTDNVKPCQVPAPEFTEVDGVLRLHYAPACKPLLRBEVTFVLGLNQYLVSQIPEPE 1140  
Db 2106 TTDNVKPCQVPAPEFTEVDGVLRLHYAPACKPLLRBEVTFVLGLNQYLVSQIPEPE 2165  
QY 1141 PDVAVLTSMLTDPDGHITAEAKRLRAGSPPSLASSASQSLASPLKATCTTRHDSPPAD 1200  
Db 2166 PDVAVLTSMLTDPDGHITAEAKRLRAGSPPSLASSASQSLASPLKATCTTRHDSPPAD 2225  
QY 1201 LIEANLLWRQBMGNITRVESKNVILDSFEPIQAEDEDEVSVPAILRRSRKFFPRAM 1260  
Db 2226 LIEANLLWRQBMGNITRVESKNVILDSFEPIQAEDEDEVSVPAILRRSRKFFPRAM 2285  
QY 1261 PIWARPDPNPLLESKWDKDPVVPVHVGCPLPPAKAPPIPPRRKRTVVLSESVSSALA 1320  
Db 2286 PIWARPDPNPLLESKWDKDPVVPVHVGCPLPPAKAPPIPPRRKRTVVLSESVSSALA 2345  
QY 1321 ELATKTFGSSSAVDSGTATASPDQPSDDGDSVESYSSMPPLEGEPDPLSDGWS 1380

Db 2346 ELATKTFGSSSAVDSGTATALPDQASDDGDSVESYSSMPPLEGEPDPLSDGWS 2405  
QY 1381 STVSEASBDDVCCMSYTWGALITPCAABETKLPINALNSLLRHNNLVYATTSAS 1440  
Db 2406 STVSEASBDDVCCMSYTWGALITPCAABETKLPINALNSLLRHNNLVYATTSAS 2465  
QY 1441 LRQKKYTFORLQVLDHRYDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
Db 2466 LRQKKYTFORLQVLDHRYDVLKEMKAKASTVAKLLSVEEACKLTPPHSARSKFGYGAK 2525  
QY 1501 DVNLSKSAVNHRISVWVKDILETETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDLG 1560  
Db 2526 DVNLSKSAVNHRISVWVKDILETETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDLG 2585  
QY 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQSPGQVFEFLYNAMKAKKCPMGAFVDTTRCFDS 1620  
Db 2586 VRVCEKMALYDVVSTLPOAVMGSSYGFQSPGQVFEFLYNAMKAKKCPMGAFVDTTRCFDS 2645  
QY 1621 TVTENDIRVEESYQCCDLAPEARQAIRSLTERLYTIGGPLTNSKGQNCYRRRCRASGVL 1680  
Db 2646 TVTENDIRVEESYQCCDLAPEARQAIRSLTERLYTIGGPLTNSKGQNCYRRRCRASGVL 2705  
QY 1681 TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLRAFTEAMTRY 1740  
Db 2706 TSCGNTLTCTYLKAAACRAAKLQDCTMLVCGDDLWICESAGTQDEASLRAFTEAMTRY 2765  
QY 1741 SAPGGPPKPEYDLELITSCSSNVVAHDASGRVYVLTDPDPTPLARAWETARHTPVN 1800  
Db 2766 SAPGGPPKPEYDLELITSCSSNVVAHDASGRVYVLTDPDPTPLARAWETARHTPVN 2825  
QY 1801 SWLGNIIIMYAPTILWARMILMTHFFSILLAOEQLKALDCQIYACYSIEPLDLPIQLORL 1860  
Db 2826 SWLGNIIIMYAPTILWARMILMTHFFSILLAOEQLKALDCQIYACYSIEPLDLPIQLORL 2885  
QY 1861 HGLSARSLHSYSPGEINRVASCLRLKGLVPLRVHRHARSVRARLLSQGGRAATCGKYLE 1920  
Db 2886 HGLSARSLHSYSPGEINRVASCLRLKGLVPLRVHRHARSVRARLLSQGGRAATCGKYLE 2945  
QY 1921 NNAVTKLKLTPIPAASQDLSSWVAGYSGGDIYHLSLRARPRFWFNCILLISVGVGIY 1980  
Db 2946 NNAVTKLKLTPIPAASQDLSSWVAGYSGGDIYHLSLRARPRFWFNCILLISVGVGIY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

## RESULT 2

A45573

genome polyprotein - hepatitis C virus (strain JT)

N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain JT) (nonstructural

protein NS4a; nonstructural protein NS4b; nonstructural protein NS5

C:Species: hepatitis C virus

C:Date: 19-May-2000 #sequence\_revision 19-May-2000 #text\_change 09-Jul-2004

C:Accession: A45573

P:Tanaka, T.; Kato, N.; Nakagawa, M.; Ootsuyama, Y.; Cho, M.J.; Nakazawa, T.; Hijikata,

Virus Res. 23, 39-53, 1992

A:Title: Molecular cloning of hepatitis C virus genome from a single Japanese carrier:

A:Reference number: A45573; MUID:92295714; PMID:1318627

A:Accession: A45573

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-3010 &lt;TAN&gt;

A:Cross-references: UNIPROT\_Q00269; GB:D11168; GB:D01171; NID:G221612; PIDN:BAA01943.1.7

A:Experimental source: HCV-JT

C:Superfamily: hepatitis C virus genome polyprotein

C:Note: sequence extracted from NCBI backbone (NCBI:106206, NCBI:106207)

C:Keywords: Atp; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; seri

F:2-115/Product: capsid protein C #status predicted &lt;CPC&gt;

F:116-191/Product: envelope protein M #status predicted &lt;EPM&gt;

F:192-389/Product: major envelope protein E #status predicted &lt;MEE&gt;

F:390-729/Product: nonstructural protein NS1 #status predicted &lt;NS1&gt;



F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F:1007-1615/Product: hepatitisin #status predicted <NS3>  
F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
F:1312-1317/Region: nucleotide-binding motif B  
F:1316-1319/Region: DEXH motif  
F:1316-1862/Product: nonstructural protein NS4a #status predicted <N4a>  
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4b>  
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match	97.11%	Score 10159	DB 1	Length 3010	Best Local Similarity 96.11%	Pred. No. 0	Mismatches 48	Conservative 29	Indels 0	Gaps 0
QY	1	MAPITAYSCQRLGCLISITGRNQVEGEVQVSTATQSFATCNGVCWTVYHGA	60							
DB	1026	LAPITAYAQQRGLGCLIVTSUTGRDKNQVEGEVQVSTATQSFATCNGVCWTVYHGA	1085							
QY	61	GSKTLAGPKGPIITQMYTNVDQDLVGWQAPPAGARSILTPCTCGSSDLYLVTRHADVIPVRRR	120							
DB	1086	GSKTLAGPKGPIITQMYTNVDQDLVGWQAPPAGARSILTPCTCGSSDLYLVTRHADVIPVRRR	1145							
QY	121	GDSRGLSPRPVSVLKGSSGGLLCPGSHAVGIFPRAAVCTRGVAKAVDFIPVESMETTM	180							
DB	1146	GDGRGLSPRPVSVLKGSSGGLLCPGSHAVGIFPRAAVCTRGVAKAVDFIPVESMETTM	1205							
QY	181	RSPVFTDNSSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAGYKVLINPSVAATLFG	240							
DB	1206	RSPVFTDNSSPPAVPQTFQVAHLHAPTGSKSTKVPAAAYAGYKVLINPSVAATLFG	1265							
QY	241	AYMSKAHGDINIRTVGRTITTGAPITYSTYKFLADGCGSGAYDIIICECHSTDSTT	300							
DB	1266	AYMSKAHGDINIRTVGRTITTGAPITYSTYKFLADGCGSGAYDIIICECHSTDSTT	1325							
QY	301	ILGIGTVLDAQETAGARLVVLATAPPGSVTVPHNIEEVALSSTGEIPFYGKAIPETI	360							
DB	1326	ILGIGTVLDAQETAGARLVVLATAPPGSVTVPHNIEEVALSSTGEIPFYGKAIPLEAI	1385							
QY	361	KGGRHLIFCHSKKKCDELAAGLGLNAYAYRGDVSIVPTSGDVIIVATDALMTGFT	420							
DB	1386	KGGRHLIFCHSKKKCDELAAGLGLNAYAYRGDVSIVPTSGDVIIVATDALMTGFT	1445							
QY	421	GDFSVIDNCVCTQVDFSLDPTTETTTVPQDAVRSORRGRTRGRMGIYRFVTPG	480							
DB	1446	GDFSVIDNCVCTQVDFSLDPTTETTTVPQDAVRSORRGRTRGRMGIYRFVTPG	1505							
QY	481	ERPSSMFDSSVLCBCEYDAGCAWYELTPAETSVRLEAYLNTPLGPVCDHLFWESVFTGL	540							
DB	1506	ERPSSMFDSSVLCBCEYDAGCAWYELTPAETSVRLEAYLNTPLGPVCDHLFWESVFTGL	1565							
QY	541	THIDAHFLSOTKQAGDNPPYLVAQATYCARAQAPPSQWQKCLIELKPTLHGPTPLL	600							
DB	1566	THIDAHFLSOTKQAGDNPPYLVAQATYCARAQAPPSQWQKCLIELKPTLHGPTPLL	1625							
QY	601	YRLGAVQNEVTTTHPIITKYMACHMSADLEWTVSTWLVGGVLAALAAAYCLTGTG	660							
DB	1626	YRLGAVQNEVTTTHPIITKYMACHMSADLEWTVSTWLVGGVLAALAAAYCLTGTG	1685							
QY	661	IILSGKPAIIPDREVLVYREFDEMEECASHLPYIEQGMOLABQFKQKATGLIQTAKQEA	720							
DB	1686	IILSGRPAVDPDREVLVYREFDEMEECASHLPYIEQGMOLABQFKQKATGLIQTAKQEA	1745							
QY	721	AAPVUESKRWTLFAFWAKHMFNFSIGIYLAGLSTLPGNPAIASLMAFTASITSLTQH	780							
DB	1746	AAPVUESKRWTLFAFWAKHMFNFSIGIYLAGLSTLPGNPAIASLMAFTASITSLTQH	1805							
QY	781	TLLFNILGCVAAQALAPPSAASAFVAGIAGAGVIGLKVLDVILAGYGAGVAGALVA	840							
DB	1806	TLLFNILGCVAAQALAPPSAASAFVAGIAGAGVIGLKVLDVILAGYGAGVAGALVA	1865							
QY	841	FKWMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVPGEGVQWNNRLIAFASR	900							
DB	1866	FKWMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVPGEGVQWNNRLIAFASR	1925							

QY	901	GNHVSPTHVPSDAAARVTOILLSSLTITOLLKRLHQMINEDCSTPCSGSWLRDWDWTC	960							
DB	1926	GNHVSPTHVPSDAAARVTOILLSSLTITOLLKRLHQMINEDCSTPCSGSWLRDWDWTC	1985							
QY	961	TVLTDKFWLQSKLLPRLPGVPPFFSCQRGYKGVWGRDGMQTTCCGQAQITGHVXNGSWR	1020							
DB	1986	TVLTDKFWLQSKLLPRLPGVPPFFSCQRGYKGVWGRDGMQTTCCGQAQITGHVXNGSWR	2045							
QY	1021	IVGPRTCNTWHGTFFPINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDHFHYTGM	1080							
DB	2046	IVGPKTCNTWHGTFFPINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDHFHYTGM	2105							
QY	1081	TTDNVKKCPQVPAPEFFTEVDGVRHLRYAPACKPLLRBEVTLVGINQYLQVCSQLCPEE	1140							
DB	2106	TTDNVKKCPQVPAPEFFTEVDGVRHLRYAPACKPLLRBEVTLVGINQYLQVCSQLCPEE	2165							
QY	1141	PDVAVLTSMLDPSHITAEAKRLRLARGSPSSLASSASQSLSAKATCTTRHDSPAD	1200							
DB	2166	PDVAVLTSMLDPSHITAEAKRLRLARGSPSSLASSASQSLSAKATCTTRHDSPAD	2225							
QY	1201	LIEANLLMRQBMGMNITRVESENKVILDSPEPQAEDEDEREVSVPAEILRRSRKFPBAM	1260							
DB	2226	LIEANLLMRQBMGMNITRVESENKVILDSPEPQAEDEDEREVSVAABILKSKKFPAL	2285							
QY	1261	PIWARPYNPILLESWKDPDYVPPVHGCPLPPAKAPDIIPPRKRTTVLSESTVSSALA	1320							
DB	2286	PIWARPYNPILLESWKSPDYVPPVHGCPLPPPTTGPPIPPRKRRTTVLSESTVSSALA	2345							
QY	1321	ELATKTFGSSSSAVDSGTATATADPDSDGSDVESYSMPPLEGEPDPLSDGWS	1380							
DB	2346	ELATKTFGSSSSAVDSGTATADPDSDGSDVESYSMPPLEGEPDPLSDGWS	2405							
QY	1381	STVSEASEDEDVCCSMSYTTWTGALITPCAAETKLPINALSNSLLRHHNLVYATTSRAS	1440							
DB	2406	STVSEASDDIVCCSMSYTTGALITPCAAEESKLPINALSNSLLRHHNMVYATTSRAS	2465							
QY	1441	LROKVTDFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKEGYGAK	1500							
DB	2466	LROKVTDFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKEGYGAK	2525							
QY	1501	DVRNLSSKAVNHRSVWKDILEDDETPTDITIMAKNEVFCVQPEKGRKPARLIVFPDLG	1560							
DB	2526	DVRNLSSKAVNHRSVWKDILEDDETPTDITIMAKNEVFCVQPEKGRKPARLIVFPDLG	2585							
QY	1561	VRCEKMAKYDVVSTLPQAVMGSSYGFQYSPQGRVEFLVNAWKAKCPMGFAYDTRCFDS	1620							
DB	2586	VRCEKMAKYDVVSTLPQAVMGSSYGFQYSPQGRVEFLVNAWKAKCPMGFAYDTRCFDS	2645							
QY	1621	TVTENDIRVESIYQCCDLAPEARQATSLTERLYIGGPLTNSKGQNGCYRRCRASGVLT	1680							
DB	2646	TVTENDIRVESIYQCCDLAPEARQATSLTERLYIGGPLTNSKGQNGCYRRCRASGVLT	2705							
QY	1681	TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLWVVCESAGTOBDEASLAFTEAMTRY	1740							
DB	2706	TSCGNLTLCYLKATACRAAKLQDCTMLVNGDDLWVVCESAGTOBDEASLAFTEAMTRY	2765							
QY	1741	SAPGDPKPKPYDELITSCSSNSVSAHDASGKRVYLTDRPTPLARAAMETARHTPVN	1800							
DB	2766	SAPGDPKPKPYDELITSCSSNSVSAHDASGKRVYLTDRPTPLARAAMETARHTPVN	2825							
QY	1801	SWLGNITMYAPTLMARMLTMTHFESILLAQOLEKALDCQIYGACYSIEPLDLPQIQLR	1860							
DB	2826	SWLGNITMYAPTLMARMLTMTHFESILLAQOLEKALDCQIYGACYSIEPLDLPQIQLR	2885							
QY	1861	HGLSAFSLHSYSGEINRVAESCLRLGVPPPLVRHRRARSVRARLLSOGGAAACGKYL	1920							
DB	2886	HGLSAFSLHSYSGEINRVAESCLRLGVPPPLVRHRRARSVRARLLSOGGAAACGKYL	2945							
QY	1921	NWAVRTKLKLTPIPAASQDLSSWFVAGYSGGDIYHSLSRAPRPFWMWCLLLSVGVGI	1980							
DB	2946	NWAVRTKLKLTPIPAASQDLSSWFVAGYSGGDIYHSLSRAPRPFWMWCLLLSVGVGI	3005							
QY	1981	LLFNR	1985							

```
Db      3006 LLENR 3010
|||||
RESULT 3
GNWVCJ
genome polyprotein - hepatitis C virus (strain J)
N:Contains: capsid protein C; envelope protein M; major envelope protein E; nonstructural
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004
C:Accession: A39253; PMID:9108950; PMID:2175903
R:Kato, N.; Hijikata, M.; Ootsuyama, Y.; Nakagawa, M.; Ohkoshi, S.; Sugimura, T.; Shimoto
Proc. Natl. Acad. Sci. U.S.A. 87, 9524-9528, 1990
A:Title: Molecular cloning of the human hepatitis C virus genome from Japanese patients
A:Reference number: A39253; MUID:9108950; PMID:2175903
A:Molecule type: genomic RNA
A:Accession: A39253
A:Residues: 1-3010 <XA2>
A:Cross-references: UNIPROT:P26662; GB:D90208; NID:g221610; PIDN:BAA14233.1; PID:g221611
R:Kato, N.; Ohkoshi, S.; Shimotohno, K.
Proc. Jpn. Acad. 65B, 219-223, 1989
A:Title: Japanese isolates of the non-A, non-B hepatitis viral genome show sequence vari
A:Reference number: PS0085
A:Accession: PS0086
A:Molecule type: genomic RNA
A:Residues: 2850-2707 <XA2>
A:Experimental source: Japanese isolate
C:Comment: The cleavage sites of this polyprotein have not been determined.
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: Atp; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:132-389/Product: major envelope protein E #status predicted <MBE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1516-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2043/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,325,417,423,430,448,532,556,576,623,645,1213,1255,2041,2077,2240,2
Query Match 96.6%; Score 10109; DB 1; Length 3010;
Best Local Similarity 95.6%; Pred. No. 0;
Matches 1896; Conservative 50; Mismatches 37; Indels 0; Gaps 0;
Qy      1 MAPITAYSOQTRGLGCIITSLTGRDNQVGEVQVYSTATQSFATCVNGVCMVTVHGA 60
Db      1026 LAPITAYSOQTRGLGCIITSLTGRDNQVGEVQVYSTATQSFATCVNGVCMVTVHGA 1095
Qy      61 GSKTLAGPKGPIQMTYNTVDQLVGMQAPPGARSITPCTCGSSDLVLTTRHADVPVRRR 120
Db      1086 GSKTLAGPKGPIQMTYNTVDQLVGMQAPPGARSITPCTCGSSDLVLTTRHADVPVRRR 1145
Qy      121 GDSRGSLLSPRPVSYLKGSGGPLLCPSHAVGIFRAVCTRGVAKAVDFVPVSMETTM 180
Db      1146 GDSRGSLLSPRPVSYLKGSGGPLLCPSHAVGIFRAVCTRGVAKAVDFVPVSMETTM 1205
Qy      181 RSPVFTDNSSPPVPPQFQVLAHLHAPTGSKGSKTKVPAAYAAQGVKVLVLPNSVAATLFGF 240
Db      1206 RSPVFTDNSSPPVPPQFQVLAHLHAPTGSKGSKTKVPAAYAAQGVKVLVLPNSVAATLFGF 1255
Qy      241 AYMSKAHGIDPNIRITGVRTITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 300
Db      1266 AYMSKAHGIBENIRITGVRTITGAPITYSTYKFLADGCGSGGAYDIIICDECHSTDSTT 1325
Qy      301 ILGIGTVLDQAEAGARLVLAATPPGTVTPHPNIEEVALSTGRIFFYKAIPIETI 360
Db      1326 ILGIGTVLDQAEAGARLVLAATPPGTVTPHPNIEEVALSTGRIFFYKAIPIETI 1385
```

```
Db 2466 LQKKVTEDRLQVLDHRYDLVKEMKAKASTVKARLSIEBACLKTPPHSAKSKFEGYCAK 2525
QY 1501 DVRLSSKAVNHRSVWKDLLEDTEPTDITIMAKNEVFCVQPEKGGKRPALRIVFPDLG 1560
Db 2526 DVRLSSKAVNHRSVWKDLLEDTEPTDITIMAKNEVFCVQPEKGGKRPALRIVFPDLG 2585
QY 1561 VRVCEKALYDVVSTLPQAVNGSSYGFQYSPQQRVEFLVNAWAKKCPMGFAIDTRCFDS 1620
Db 2586 VRVCEKALYDVVSTLPQAVNGSPSYGFQYSPQQRVEFLVNTWKSCKPMGFSYDTRCFDS 2645
QY 1621 TVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKQNGCYRRCRASGVLT 1680
Db 2646 TVTENDIRTESIYQCCDLAPEARQAIRSLTERLYVGGPLTNSKQNGCYRRCRASGVLT 2705
QY 1681 TSCGNLTTCYLKAAACRAAKLQDCTMLVCGDDLNVVICESAGTQDEASLRAFTAMTRY 1740
Db 2706 TSCGNLTTCYLKATACRAAKLQDCTMLVNGDDLNVVICESAGTQDEAALRAFTAMTRY 2765
QY 1741 SAPGDPKPEYDLELITSCSNVSVVAHDASGRVYVYLTRDPTTPARAAWETARHTPVN 1800
Db 2766 SAPGDPKPEYDLELITSCSNVSVVAHDASGRVYVYLTRDPTTPARAAWETARHTPVN 2825
QY 1801 SWLGNIMYAPTLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPQIOLR 1860
Db 2826 SWLGNIMYAPTLWARMILMTHFFSILLAOBLEKALDCQIYGACYSIEPLDLPQIOLR 2885
QY 1861 HGLSAFSLHSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPG 1920
Db 2886 HGLSAFSLHSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPGSPG 2945
QY 1921 NWAVENTKULPTIPAASOLDLSSWVAGYSGGDIYHSLSRAPRPFWMCLLLSVGVGIY 1980
Db 2946 NWAVENTKULPTIPAASOLDLSSWVAGYSGGDIYHSLSRAPRPFWMCLLLSVGVGIY 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010

RESULT 4
S18030
genome polyprotein - hepatitis C virus (isolate JK1)
N:contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructu
C:Species: hepatitis C virus
A:Variety: isolate JK1
C:Date: 19-May-2000 #sequence revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S18030; S33570; A48332; S18029
R:Honda, M.; Kaneko, S.; Masashi, U.; Kobayashi, K.; Murakami, S.
submitted to the EMBL Data Library, September 1991
A:Description: A whole genome of hepatitis C virus cDNA was isolated from a single patie
A:Reference number: S18028
A:Accession: S18030
A:Molecule type: genomic RNA
A:Residues: 1-3010 <HON>
A:Cross-references: UNIPROT:Q68949; EMBL:X61596; NID:G59478; PIDN:CAAA3793.1; PID:G59479
A:Experimental source: isolate JK1 from an individual
R:Honda, M.; Kaneko, S.; Unoura, M.; Kobayashi, K.; Murakami, S.
Arch. Virol. 128, 163-169, 1993
A:Title: Sequence analysis of putative structural regions of hepatitis C virus isolated
A:Reference number: A48332; MUID:93119270; PMID:8380322
A:Accession: S33570
A:Molecule type: genomic RNA
A:Residues: 1-547, 'T', 549-621, 'V', 623-624, 'S', 626-652, 'DL', 655-761, 'T', 763-782 <HON>
A:Cross-references: EMBL:X61591
A:Note: this sequence is inconsistent with the nucleotide translation
A:Note: the authors translated the codon AGG for residue 43 as Pro, TGG for residue 320
as Trp, and TTC for residue 771 as Ser
A:Note: sequence extracted from NCBI backbone (NCBI:121747, NCBI:121748)
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
```

```
F:116-131/Product: envelope protein M #status predicted <EPM>
F:192-389/Product: major envelope protein E #status predicted <MEE>
F:730-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:1007-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1230-1615/Product: hepatitis C virus #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2010/Product: nonstructural protein NS5 #status predicted <NS5>
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>
F:196,209,234,250,305,417,423,448,532,540,556,576,623,645/Binding site: carbohydrate (As
```

```
Query Match 96.2%; Score 10069; DB 1; Length 3010;
Best Local Similarity 95.8%; Pred. NO. 0;
Matches 1901; Conservative 39; Mismatches 45; Indels 0; Gaps 0;

QY 1 MAPITAYSOOTRGLGGLTITSLTGRDRNOVGEVQVSTATQSFATCVNGVCMVYHGA 60
Db 1026 LAPITAYSOOTRGLGGLTITSLTGRDRNOVGEVQVSTATQSFATCVNGVCMVYHGA 1085
QY 61 GSKTLAGEPKGITOMYTNVDQDLVGMQAPPGARSITPCTCGSSDLYLVTRHADVLPVRR 120
Db 1086 GSKTLAGEPKGIPINQMYTNVDQDLVGMQAPSGAASITPCTYSSDLYLVTRHADVLPVRR 1145
QY 121 GDSRGSLLSPRPVSVYKSGSGGPLLCPGSHAVGIFPRAAVCTRGVAKAVDFVPVSEMETTM 180
Db 1146 GDSRGSLLSPRPVSVYKSGSGGPLLCPGSHAVGIFPRAAVCTRGVAKAVDFIPVSEMETTM 1205
QY 181 RSPVFTDNSSPPAVPQTEFQVAHLHAPTGSKSTKVPAAVAAQGYKVLNPSVAATLFG 240
Db 1206 RSPVFTDNSSPPAVPQTEFQVAHLHAPTGSKSTKVPAAVAAQGYKVLNPSVAATLFG 1265
QY 241 AYMSKAHGDINIRITGVRTITTGAPITVYSTYKFLADGCGSGGAYDIIICDECHSDST 300
Db 1266 AYMSKAHGDINIRITGVRTITTGAPITVYSTYKFLADGCGSGGAYDIIICDECHSDST 1325
QY 301 ILGIGTVLDQAEATAGARLVVLAATPPGSVTVPHNIEEVALSSTGETIPFYKALPIETI 360
Db 1326 ILGIGTVLDQAEATAGARLVVLAATPPGSVTVPHNIEEVALSSTGETIPFYKALPIETI 1385
QY 361 KGRHLIFCHSKKKDELAALKSLGLNVAAYVGLDVSIVPTSGDVLVVAATDALMTGFT 420
Db 1386 KGRHLIFCHSKKKDELAALKSLGLNVAAYVGLDVSIVPTSGDVLVVAATDALMTGFT 1445
QY 421 GDFSVIDCNTCVITQVDFSLDPTFTIETTTVPQDAVSRQRGRGTGRGRGIYFVTPG 480
Db 1446 GDFSVIDCNTCVITQVDFSLDPTFTIETTTVPQDAVSRQRGRGTGRGRGIYFVTPG 1505
QY 481 ERPSGMFSSVLCBCYDAGCAWYELTPAETSVRLRAYLNTFGLPVCQDHLFWESVFTGL 540
Db 1506 ERPSGMFSSVLCBCYDAGCAWYELTPAETSVRLRAYLNTFGLPVCQDHLFWESVFTGL 1565
QY 541 THIDAHFLSOTKOAGDNPPYLVAQVATCARAQAPPPSDMDQMKCLIRLKLPTLHGTPLL 600
Db 1566 THIDAHFLSOTKOAGDNPPYLVAQVATCARAQAPPPSDMDQMKCLIRLKLPTLHGTPLL 1625
QY 601 YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTSVVIYGR 660
Db 1626 YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGTSVVIYGR 1685
QY 661 IILSKPAIIPDREVLVREFFEMEECASHLPIYBQGMOLARQFKQKALGLLTATKQAE 720
Db 1686 IILSKPAIIPDREVLVREFFEMEECASHLPIYBQGMOLARQFKQKALGLLTATKQAE 1745
QY 721 AAPVVSQWRTLEAFWAKHWNFTSGIYLAGLSTLPCNPALASIMAFSTSTPLTTOH 780
Db 1746 AAPVVSQWRTLEAFWAKHWNFTSGIYLAGLSTLPCNPALASIMAFSTSTPLTTOH 1805
QY 781 TLLFNILGWSVAALAPPASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVAGALVA 840
Db 1806 TLLFNILGWSVAALAPPASAFVAGIAGAAVSGISGLGKVLVDILAGYGAGVAGALVA 1865
```

QY 841 FKVMGEMPTEDLVNLLPALLSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAPSR 900  
Db 1866 FKMGSEMPSTEDLVNLLPALLSPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAPSR 1925  
QY 901 GNHVSPTHYVPESDAARVTKILSSLLITQLKRLHWINEDCSTPCSGSLRDLVMDWIC 960  
Db 1926 GNHVSPTHYVPESDAARVTKILSSLLITQLKRLHWINEDCSTPCSGSLRDLVMDWIC 1985  
QY 961 TVLTDFTXWLSKLLPLPGVPPFPSCQGRYKGVWRGDIQMTTCPCGAQITGHVKGSMR 1020  
Db 1986 TVLTDFTXWLSKLLPLPGVPPFPSCQGRYKGVWRGDIQMTTCPCGAQITGHVKGSMR 2045  
QY 1021 IVPBRTCSNTHGTFPFINAVTGTCTSPAPNYSRALWRVAAEYVSVTRVGDHYVTGM 1080  
Db 2046 IVPBRTCSNTHGTFPFINAVTGTCTSPAPNYSRALWRVAAEYVSVTRVGDHYVTGM 2105  
QY 1081 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPLLRDEEYFVLGNQLVLSGSLPCEPE 1140  
Db 2106 TTDNVKCPQVPAPEFFTEVDGVLHRYAPACKPLLRDEEYFVLGNQLVLSGSLPCEPE 2165  
QY 1141 PDVAVLTSMLTDSHITAEAKRLARGSPSLASSASQLSAPSLKATCTTRHDSPAD 1200  
Db 2166 PDVTVLTSMLTDSHITAEAKRLARGSPSLASSASQLSAPSLKATCTTRHDSPAD 2225  
QY 1201 LIEANLLWROEMGMGNITRVESENKVVILDSFEPLOAEDEEREVGPVPAEILRRSKFPFRAM 1260  
Db 2226 LIEANLLWROEMGMGNITRVESENKVVILDSFEPLOAEDEEREVGPVPAEILRRSKFPFRAM 2285  
QY 1261 PIWARPDPNPLLESWKDDPYVPVPHVGCPLPPAKAPPPIPPRRKRTVWLSESTVSALA 1320  
Db 2286 PIWARPDPNPLLESWKDDPYVPVPHVGCPLPPAKAPPPIPPRRKRTVWLSESTVSALA 2345  
QY 1321 ELATKTGSESSAVDSGTATASDQSDDDGADSDVESYSSMPLPGEFGDPLDLSGSW 1380  
Db 2346 ELATKTGSESSAVDSGTATASDQSDDDGADSDVESYSSMPLPGEFGDPLDLSGSW 2405  
QY 1381 STVSEERASEDVCCSMSTYTGALITPCAAEETKLPINALSNSLLRHHNLVYATTSAS 1440  
Db 2406 STVSEERASEDVCCSMSTYTGALITPCAAEESKLPINLSNLSLLRHHNVYATTSAS 2465  
QY 1441 LRQKVTFRDRLQVLDHYRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSFGYGA 1500  
Db 2466 LRQKVTFRDRLQVLDHYRDVLKEMKAKASTVAKALLSVEEACKLTPPHSARSFGYGA 2525  
QY 1501 DVRLSSKANHRSVMKOLLEDETETIDITIMAKNEVFCVQPEKGRKPARLTVFPDLG 1560  
Db 2526 DVRLSSKANHRSVMKOLLEDETETIDITIMAKNEVFCVQPEKGRKPARLTVFPDLG 2585  
QY 1561 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQVFEFLVNAWKAKCPMGFAYDTRCFDS 1620  
Db 2586 VRVCEKMALYDVVSTLPQAVMGSSYGFQYSPGQVFEFLVNAWKAKCPMGFAYDTRCFDS 2645  
QY 1621 TVTENDIRVEESYQCCDLAPARQARSUTERLYIGGLPTNSKGQCYRRRCRASGLT 1680  
Db 2646 TVTESDIRVEESYQCCDLAPARQARSUTERLYIGGLPTNSKGQCYRRRCRASGLT 2705  
QY 1681 TSCGNTLTCYKAAACRAAKLOCTMLVCGDDLAVLICESAGTCODEASRAFTTEATRY 1740  
Db 2706 TNCGNTLTCYKAAACRAAKLOCTMLVCGDDLAVLICESAGTCODEASRAFTTEATRY 2765  
QY 1741 SAPPGDPKPEYDLELITSCSNVSVADHDSAGKRVYVLTDRPTTPLARAWEATRHPTVN 1800  
Db 2766 SAPPGDPKPEYDLELITSCSNVSVADHDSAGKRVYVLTDRPTTPLARAWEATRHPTVN 2825  
QY 1801 SWLGNITMYAPTLLWARMILTHFSSILLAQOLEKALDCCQIYGACYSIEPLDLPQIITQRL 1860  
Db 2826 SWLGNITMYAPTLLWARMILTHFSSILLAQOLEKALDCCQIYGATYIEPLDLPQIITQRL 2885  
QY 1861 HGLSAPSLHSYSGEINRVASCLRLKGVPEPLVRWRHARSVRARLLSOGGRAATCGKYL 1920  
Db 2886 HGLSAPSLHSYSGEINRVASCLRLKGVPEPLVRWRHARSVRARLLSOGGRAATCGKYL 2945  
QY 1921 NNAVRTKLTPTPAASQLDLSWFWVAGYSGGDIYHLSLRARPRFWFMCCLLLSVGVGIY 1980

Db 2946 NNAVRTKLTPTPAASQLDLSWFWVAGYSGGDIYHLSLRARPRFWFMCCLLLSVGVGIY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010  
RESULT 5  
GNVWTW  
genome polyprotein - hepatitis C virus (strain Taiwan)  
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain Taiwan)  
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5  
C:Species: hepatitis C virus  
A:Note: host Homo sapiens (man)  
C:Date: 31-Dec-1992 #sequence\_revision 31-Dec-1992 #text\_change 09-Jul-2004  
C:Accession: A40244  
R:Chen, P.J.; Lin, M.H.; Tai, K.F.; Liu, P.C.; Lin, C.J.; Chen, D.S.  
Virology 188, 102-113, 1992  
A:Title: The Taiwanese hepatitis C virus genome: sequence determination and mapping the  
A:Reference number: A40244; MUID:92230206; PMID:1314449  
A:Accession: A40244  
A:Molecule type: genomic RNA  
A:Residues: 1-3010 <CHE>  
A:Cross-references: UNIPROT:P29846; GB:M84754  
C:Superfamily: hepatitis C virus genome polyprotein  
C:Keywords: Atp; capsid protein; envelope protein; glycoprotein; hydrolase; nonstructur  
F:1-115/Product: capsid protein C #status predicted <CPC>  
F:116-191/Product: envelope protein M #status predicted <EPM>  
F:192-389/Product: major envelope protein E #status predicted <MEE>  
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>  
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>  
F:1007-1615/Product: hepatitis C virus genome polyprotein NS2 #status predicted <NS2>  
F:1230-1237/Region: nucleotide-binding motif A (P-loop)  
F:1312-1317/Region: nucleotide-binding motif A (P-loop)  
F:1316-1319/Region: DEXH motif  
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>  
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>  
F:2014-3010/Product: nonstructural protein NS5 #status predicted <NS5>  
F:196,209,233,234,250,325,417,423,430,448,532,540,556,576,623,1213,1255,2041,20  
Query Match 96.2%; Score 10066; DB 1; Length 3010;  
Best Local Similarity 95.4%; Pred. No. 0;  
Matches 1894; Conservative 44; Mismatches 47; Indels 0; Gaps 0;  
QY 1 MAPIYASQOTRGLGCIITSLTRGRDNQVEGEVQVVSATQSFATCNGVCWTVYHGA 60  
Db 1026 LAPITAYAQOTRGLGCIITSLTRGRDNQVEGEVQVVSATQSFATCNGVCWTVYHGA 1085  
QY 61 GSKTLAGPKGITOMYTNVDQDLVGMQAPPGASLTCTCGSSDLYLVRHADVIPVRR 120  
Db 1086 GSKTLAGPKGITOMYTNVDQDLVGMQAPPGASLTCTCGSSDLYLVRHADVIPVRR 1145  
QY 121 GDSRGSLLSPRVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDVPVSMETTM 180  
Db 1146 GDSRGSLLSPRVSYLKGSSGGPLLCPSGHVVGIFRAAVCTRGVAKAVDVPVSMETTM 1205  
QY 181 RSPVFTDNSPFAVQPTFOVAHLHAPTGSCKSTKVPAAAYAAQGYKVLNPSVAATLFG 240  
Db 1206 RSPVFTDNSPFAVQPTFOVAHLHAPTGSCKSTKVPAAAYAAQGYKVLNPSVAATLFG 1265  
QY 241 AYMSKAGIDPNIRTVRTITTCAPTISTYVGFADGGCGGAYDIIICDECHSTDTT 300  
Db 1266 AYMSKAGIDPNIRTVRTITTCAPTISTYVGFADGGCGGAYDIIICDECHSTDTT 1325  
QY 301 ILGIGTVLDOAETAGARLVVLAATATPGSVTVPHNPTEEVALSTGIPFGKAIPIETI 360  
Db 1326 ILGIGTVLDOAETAGARLVVLAATATPGSVTVPHNPTEEVALSTGIPFGKAIPIETI 1385  
QY 361 KGRHLLIFCHSKKCDLAAKLSGLGNNAVYRGLDVSIVPTSGDVIIVATDALMTGFT 420  
Db 1386 KGRHLLIFCHSKKCDLAAKLSGLGNNAVYRGLDVSIVPTSGDVIIVATDALMTGFT 1445

Qy	421	GDPSVIDDNCVTQVTQVDSLDFTFIETTTVPQDAVSRSQRGRGTGRGMGLYRFVTPG	480
Db	1446	GDPSVIDDNCVTQVTQVDSLDFTFIETTTMPQDAVSRSQRGRTSRGRGTYRREVTPG	1505
Qy	481	ERPSGMFDSVLCCEYDAGCAWVELTPAETSUVRLAYLNTPGUPVQDHLFEWESVFTGL	540
Db	1506	ERPSGMFDSVLCCEYDAGCAWVELTPAETSUVRLAYLNTPGUPVQDHLFEWESVFTGL	1565
Qy	541	THIDAFLSQTQAGDNFPYVAYQATVCARAQAPPSPMDQMKCLIRLKPTLHGPTPLL	600
Db	1566	THIDAFLSQTQAGDNFPYVAYQATVCARAQAPPSPMDQMKCLTRLKPTLHGPTPLL	1625
Qy	601	YRLGAVQNEVTTHTPIKTYIMACMSADLEWVTSTWLVGGVLAALAAAYCLITGCVVIVGR	660
Db	1626	YRLGAVQNEVTLTHPIKTYIMACMSADLEWVTSTWLVGGVLAALAAAYCLITGCVVIVGR	1685
Qy	661	IIILSGPAILPDREVLYRFEDEMEECASHLPYTEQGMQLAEQPKOKAIGLQATKQAEA	720
Db	1686	IIILSGPAPVDPREVLYQFEDEMEECASHLPYTEQGMQLAEQPKQALGULQATKQAEA	1745
Qy	721	AAPVWESKWRTLBAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLAFTASITSPLTQH	780
Db	1746	AAPVWESKWRTLBAFWANDMWNFISGIQYLAGLSTLPGNPAIASLAFTASITSPLTQS	1805
Qy	781	TLFLNIIIGHWAAQLAPPSAASFVPGAGIAGAAGVSGIGLKVLDIILAGVGAGVALVA	840
Db	1806	TLFLNIIIGHWAAQLAPPGAASAFVPGAGIAGAAGVSGIGLKVLDVMVAGYGAGVALVA	1865
Qy	841	FKVMSGEMPSTEDLVNLLPAILSPGALVGVCAAILRRHVGPBGAVQWNNRLIAPASR	900
Db	1866	FKVMSGEMPSTEDLVNLLPAILSPGALVGVCAAILRRHVDPGEGAVQWNNRLIAPASR	1925
Qy	901	GNHVSPTHYPVPSDAAARVTQILSSLTIITQLKRLHOWINEDCSTPCSGSMLRDVNDWIC	960
Db	1926	GNHVSPTHYPVPSDAAARVTQILSGLTIITQLRRLHOWINEDCSTPCSGSMLRDVNDWIC	1985
Qy	961	TVLTDFTKWLQSKLLPRLPGVPFFSCQRYKGVWRGDGIMQTTCPGQAITGHVKGSGMR	1020
Db	1986	TVLADFKTILQSKLLPRLPGVPFFSCQRYKGVWRGDGIMQTTCPGQAITGHVKGSGMR	2045
Qy	1021	IVGPRICSNTHGTFPFINAYTTGCPITSPAPNYSRALWRVAEEYVEVTVGDFHYVTGM	1080
Db	2046	IVGPKTCSTNTHGTFPFINAYTTGCTPSAPNYSRALWRVAEEYVEVRVGVDFHYVTGM	2105
Qy	1081	TTDNVCKPCQVPAPEFTEVDGVLRLHRYAPACKPLLEEYTFVLGLNQYLWGSOLPCEPE	1140
Db	2106	TTDNVCKPCQVPAPEFTEVDGVLRLHRYAPACKPLLEEYSFQVLNQYVVGSQLPCEPE	2165
Qy	1141	PDVAVLTSMLTDPESHITAETAKRLARGSPPSLIASSASOLSPSLKATCTTTRHDSDDAD	1200
Db	2166	PDVAVLTSMLTDPESHITAETAKRLARGSPPSLIASSASOLSLSLXAACTTTRHTPPDAD	2225
Qy	1201	LIEANILLWROEMGNNITRVESENKVTILDSFEPLQAEDEREVSVPAEILRRSRKFFPRAM	1260
Db	2226	LIEANILLWROEMGNNITRVESENKVTILDSFDPLRAEEDREVSVPAEILLRKSRKFFPAL	2285
Qy	1261	PIWARPDYNPPLLESWKDPDYPVPVHGCPPLPPAKAPPIPPRRKRTVWLSESTVSSALA	1320
Db	2286	PVWARPDYNPPLLEPKWDPDYPVPVHGCPPLPVKAPPIPPRRKRTVWLTESTVSSALA	2345
Qy	1321	ELATKTFGSSSESAVDSGTATASDPQSDGSDGSDVESYSSMPPEGEQPDPLSDGWSW	1380
Db	2346	ELATKTFGSSSESAAGSGTATAPDQPSDDGSDGSDVESCSMPPEGEQPDPLSDGWSW	2405
Qy	1381	STVSEESASEDVVCCSMGYTWTGALITPCAABETKLPINALSNSLLRHNLYVATTSRSAS	1440
Db	2406	STVSEEDGEVIGCCSMGYTWTGALITPCAABESKLPINALSNSLLRHNVMYATTSRSAS	2465
Qy	1441	LROKKVTFDRLOVLDDHYRDLVKEMKAKASTVKVAKLLSVEACKLTPPHSARSKFGYGAK	1500
Db	2466	QRQKKVTDRLQVLDDHYRDLVKEMKAKASTVKVAKLLSVEACKLTPPHSARSKFGYGAK	2525
Qy	1501	DVRNLSSKANVNHRSVMKDLLEDTEDETIDTTIMAKNEVFCVQPKGGRKPARLIVFPDLG	1560

Db	2526	DVRNLGSKAHNHSRVWVKDLEETETPDITIIWAKNEVFCVQPEKGRKPARLIVFFDLG	2589
Qy	1561	VRVCEKMALDYVYSTLPQAVMGSSYGFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS	1620
Db	2586	VRVCEKMALDYVYSTLPQAVMGSSYGFQYSPGQRFVFLVNAWAKKCPMGFAYDTRCFDS	2645
Qy	1621	TVTENDIRVEESIIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVL	1680
Db	2646	TVTENDIRVEESIIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCGYRRCRASGVL	2705
Qy	1681	TSCGNTLTCVLKAAACRAAKLOCTMLVCGDDLWVICSAGTQEDASLRATFTAMTRY	1740
Db	2706	TSCGNTLTCVLKAAACRAAKLOCTMLVCGDDLWVICSAGTQEDASLRATFTAMTRY	2765
Qy	1741	SAPPDGPPEKPEYDLELITSCSSNVSAHDASGRKRVYVLTDPDTPPLARAAWETARHTFPV	1800
Db	2766	SAPPDGLPQPEYDEQLITSCSSNVSAHDASGRKRVYVLTDPDTPPLARAAWETARHTFPV	2825
Qy	1801	SWLGNITMYAPTLMWARMILMTHFFSILLAOEQLKALDCOIQYACYSIEPDLPOIIRL	1860
Db	2826	SWLGNITMYAPTLMWARMILMTHFFSILLAOEQLKALDCOIQYACYSIEPDLPOIIRL	2885
Qy	1861	HGLSAPFSLHSYSGEINRVASCRLKGVPLRVWHRARSVRALLSQGGRAATCGKYL	1920
Db	2886	HGLSAPFSLHSYSGEINRVASCRLKGVPLRVWHRARSVRALLSQGGRAATCGKYL	2945
Qy	1921	NWAVRTKLTLPAAASQDLSSWFSVAGYSGGGDIYHSLSRARPRFWFMCLLLSLVGVGI	1980
Db	2946	NWAVRTKLTLPAAASQDLSSWFSVAGYSGGGDIYHSLSRARPRFWFMCLLLSLVGVGI	3005
Qy	1981	LLPNR 1985	
Db	3006	LLPNR 3010	

RESULT 6

GNMW33

Genome polyprotein - hepatitis C virus (strain HCV-1)

N:Contains: capsid protein C; envelope protein M; hepatitis C virus (strain HCV-1)

C:Species: hepatitis C virus

C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004

C:Accession: A39166; PQ0403; PQ0404

R:Choo, Q.L.; Richman, K.H.; Han, J.H.; Berger, K.; Lee, C.; Dong, C.; Gallegos, Proc. Natl. Acad. Sci. U.S.A. 88, 2451-2455, 1991

A:Title: Genetic organization and diversity of the hepatitis C virus.

A:Reference number: A39166; MUID:91172826; PMID:1848704

A:Accession: A39166

A:Molecule type: mRNA

A:Residues: 1-3011 <CHO>

A:Cross-references: UNIPROT:p26664; GB:M62321; NID:g329873; PIDN:AAA45676.1; P:Chan, S.W.; McOmish, F.; Holmes, E.C.; Dow, B.; Peutherer, J.F.; Follett, E.N.; J. Gen. Virol. 73, 1131-1141, 1992

A:Title: Analysis of a new hepatitis C virus type and its phylogenetic relationship with other hepatitis C virus types and their putative progenitors

A:Reference number: PQ0393; MUID:92268871; PMID:1316939

A:Accession: PQ0403

A:Molecule type: genomic RNA

A:Residues: 1577-1633 <CHA>

A:Cross-references: DDBJ:D10128

A:Experimental source: isolates E-b16

A:Accession: PQ0404

A>Status: preliminary

A:Molecule type: genomic RNA

A:Residues: 1577-1633 <CH2>

A:Experimental source: isolates E-b17

C:Superfamily: hepatitis C virus genome polyprotein

C:Keywords: ATP; capsid protein; envelope protein; glycoprotein; hydrolase; non-structural protein; product; capsid protein C #status predicted <CPC>

F:115-191/Product: capsid protein M #status predicted <CPM>

F:192-389/Product: major envelope protein E #status predicted <MEE>

F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>

F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>

[illegible][illegible]

QY 1979 TYLLPNR 1985



```

Db      3005 IYLLPNR 3011
|||||
RESULT 7
S40770
genome polyprotein - hepatitis C virus
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural
protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: S40770; PC1285
R:Okamoto, H.
submitted to the EMBL Data Library, March 1992
A:Reference number: S40770
A:Accession: S40770
A:Molecule type: genomic RNA
A:Residues: 1-3011 <OKA>
A:Cross-references: UNIPROT:Q03463; EMBL:D10749; NID:g221586; PIDN:BAA01582.1; PID:g2215
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Yotsumoto, S.; Tanaka, T.; Yoshizawa, H.; Tsuda,
Jpn. J. Exp. Med. 60, 167-177, 1990
A:Title: The 5'-terminal sequence of the hepatitis C virus genome.
A:Reference number: PC1284; MUID:91013116; PMID:2170712
A:Accession: PC1285
A:Molecule type: genomic RNA
A:Residues: 1-513 <OK2>
A:Cross-references: GB:D00831; NID:g221511; PIDN:BAA00705.1; PID:g221512
A:Experimental source: isolate HC-J1
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: Antp; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serin
F:2-115/Product: capsid protein C #status predicted <CPC>
F:116-191/Product: envelope protein M #status predicted <BPM>
F:192-389/Product: major envelope protein E #status predicted <MBE>
F:390-729/Product: nonstructural protein NS1 #status predicted <NS1>
F:730-1006/Product: nonstructural protein NS2 #status predicted <NS2>
F:1007-1615/Product: hepatitis C virus #status predicted <NS3>
F:1230-1237/Region: nucleotide-binding motif A (P-loop)
F:1312-1317/Region: nucleotide-binding motif B
F:1316-1319/Region: DEXH motif
F:1616-1862/Product: nonstructural protein NS4a #status predicted <N4A>
F:1863-2013/Product: nonstructural protein NS4b #status predicted <N4B>
F:2014-3011/Product: nonstructural protein NS5 #status predicted <NS5>

Query Match      89.8%; Score 9398.5; DB 1; Length 3011;
Best Local Similarity 87.6%; Pred. No. 0;
Matches 1740; Conservative 132; Mismatches 112; Indels 3; Gaps 3;

Qy      1 MAPITAYSQOTRGLGCIITSLTGRDRNOVEGEVQVSTQSFATCVCNGVCMVYHGA 60
Db      1026 LAPITATQQRGLGCIITSLTGRDRKNQVEGEVQVSTAAQTFLATCVCNGVCMVYHGA 1085

Qy      61 GSKTLAGPKGITOMYTNVDQDLVGWQAPPGARSILTCTCGSSDLVLTTRHADYIPVRRR 120
Db      1086 GTRTIASPKGPIQMYTNVDQDLVGWQAPPGARSILTCTCGSSDLVLTTRHADYIPVRRR 1145

Qy      121 GDSRGLSLRPVSYLKGSSGGPILCPGSHAVGIFRAVCTRGVAKAVDFVPVSMETTM 180
Db      1146 GDSRGLSLRPISYLKGSSGGPILCPAGHVVGIFRAVCTRGVAKAVDFIPVPSLETMM 1205

Qy      181 RSPVFTNNSPPVQTFQVAHLHAPTSGSKSTKVPAAAYAGQYKVLVLPNSVAATLFGF 240
Db      1206 RSPVFTNNSPPVQTFQVAHLHAPTSGSKSTKVPAAAYAGQYKVLVLPNSVAATLFGF 1265

Qy      241 AYMSKAHGIDNIRTVRTITGAPITYSTYKGLADGCGSGGAYDIIICDECHSTDSTT 300
Db      1266 AYMSKAHGIDNIRTVRTITGSPITYSTYKGLADGCGSGGAYDIIICDECHSTDATS 1325

Qy      301 ILGTGTVLDAQETAGARLVILATATPPGSVTVPHNIEVALSSTGEIPFYGKAIPETI 360
Db      1326 VLGTGTVLDAQETAGARLVILATATPPGSIIVPHNIEVALSSTGEIPFYGKAIPLEAI 1385

Qy      361 KGGRHILFCHSKKKDELAALKLGLGNVAYYRGLDVSVIPTSGLDVIVVATDALMTGFT 420

```

```

Db      1386 KGGRHILFCHSKKKDELAALKLGLGNVAYYRGLDVSVIPTSGLDVIVVATDALMTGFT 1445
Qy      421 GFDSVIDCNTCVTQVDFSLDPTFTTITVPODAVSRQRRTGRGMGIIVRFVTPG 480
Db      1446 GFDSVIDCNTCVTQVDFSLDPTFTTITVPODAVSRQRRTGRGMGIIVRFVAPG 1505
Qy      481 ERSGMFDSSVLCEDYDAGCANTVELTPEATSVRLRAYLNTPLPVCODHLEFMSVSTGL 540
Db      1506 ERSGMFDSSVLCEDYDAGCANTVELTPEATSVRLRAYLNTPLPVCODHLEFMSVSTGL 1565
Qy      541 THIDAFLSQTQAGDNFFYLVAQATVCARAQAPPPSWQMMKCLIRLKLPTLHGPTPL 600
Db      1566 THIDAFLSQTQAGDNFFYLVAQATVCARAQAPPPSWQMMKCLIRLKLPTLHGPTPL 1625
Qy      601 YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLITGWSVIVGR 660
Db      1626 YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLITGWSVIVGR 1685
Qy      661 IILSGKPAIIPDREVLRYREDEMEECASHLPYIEQGMQLAEQKQKALGILLOTATKQAEA 720
Db      1686 IIVSGRPAIIPDREVLRYREDEMEECASHLPYIEQGMQLAEQKQKALGILLOTATKQAEA 1745
Qy      721 AAPVRESKRTLBAFWAKHWMNFISGIQYLAGLSTLPGNPAIASMAFTASITSPITQOH 780
Db      1746 IAPVQINWQKLEAFWAKHWMNFISGIQYLAGLSTLPGNPAIASMAFTASITSPITQOH 1805
Qy      781 TLLFNILGGVAAQLAPPSSAASAFVAGIAGAAVSGISGLKVLVDIILAGYAGVAGALVA 840
Db      1806 TLLFNILGGVAAQLAPPSSAASAFVAGIAGAAVSGISGLKVLVDIILAGYAGVAGALVA 1865
Qy      841 FKVMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEVAVQWNNRIAPASR 900
Db      1866 FKVMSGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGEVAVQWNNRIAPASR 1925
Qy      901 GNHVSPPHYVPESDAAARVTQILSSLTITQLLKHQWINEDECTPCSGSLRWLDWDMIC 960
Db      1926 GNHVSPPHYVPESDAAARVTQILSSLTITQLLKHQWINEDECTPCSGSLRWLDWDMIC 1985
Qy      961 TVLTDFKTLQSKLLPRLPGVPFSCQYKGVWVGIMQTTCCGQAOITGHVKNKSMR 1020
Db      1986 EVLSDFKTLQSKLLPRLPGVPFSCQYKGVWVGIMQTTCCGQAOITGHVKNKSMR 2045
Qy      1021 IVPRTCSNTHGTFPPINAVTTGCTPSAPNYSRALWRVAAAEVYEVTRVGDHYHTGM 1080
Db      2046 IVPRTCSNTHGTFPPINAVTTGCTPSAPNYSRALWRVAAAEVYEVTRVGDHYHTGM 2105
Qy      1081 TTDNVKCPQVPAPEFFTEVDGVLRLHRYAPACKPLLEEVTFVLGNQYLVGSLPCEPE 1140
Db      2106 TTDNVKCPQVPAPEFFTEVDGVLRLHRYAPACKPLLEEVTFVLGNQYLVGSLPCEPE 2165
Qy      1141 PDVAVLTSMLTDPDSHITAEAKRLARGSPPSLAASSASQSLAPSLKATCTTRHDSDDAD 1200
Db      2166 PDVAVLTSMLTDPDSHITAEAKRLARGSPPSLAASSASQSLAPSLKATCTTRHDSDDAD 2225
Qy      1201 LIEANLLWRQEMGNITRVSEKNNVILDSFELPQAEDEEREVSVPALILRRSKFEPAM 1260
Db      2226 LIEANLLWRQEMGNITRVSEKNNVILDSFELPQAEDEEREVSVPALILRRSKFEPAM 2285
Qy      1261 PIWAPDYNPLLESWKDPDYPVPVHGCPLPAKAPPIPPRRKRVTWLSESTVSSALA 1320
Db      2286 PIWAPDYNPLLESWKDPDYPVPVHGCPLPAKAPPIPPRRKRVTWLSESTVSSALA 2345
Qy      1321 ELATKTFGSSSESSAVDSGTATASDPQSDDG - DAGSDVESYSSMPPLEGPGDPLSDGS 1379
Db      2346 ELATKTFGSSSESSAVDSGTATASDPQSDDG - DAGSDVESYSSMPPLEGPGDPLSDGS 2404
Qy      1380 WSTVSEEA - SEDVVVCCSMSTWTGALITPCAABETKLPINALSNILRHNLVYATTSRS 1438
Db      2405 WSTVSEEA - SEDVVVCCSMSTWTGALITPCAABETKLPINALSNILRHNLVYATTSRS 2464
Qy      1439 ASLRQKVVTFDRLOVLDHVRDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFGY 1498
Db      2465 ASLRQKVVTFDRLOVLDHVRDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFGY 2524

```



```
Db      1986 EVLSDFKTLKAKLMPQLGPIFFVSCQGRGVRWGRGIMHTRCHCAEITGHVKNQTM 2045
Qy      1021 IVGPRTCSNTHGTFPFINAYTTGCTPSPAPNSRALRWAAAEYEVVTRVGDPHYVTGM 1080
Db      2046 IVGPRTCKNWSGTFFINAYTTGCTPLPAPNYKFALRWVAEYVIRRVGDPHYVSGM 2105
Qy      1081 TTDNVKCPQVPAPPEFTTEVDGVRHLHYAPACKPLLRBEVTVFVGLNQYLVGSLPCEPE 1140
Db      2106 TTDNLKCPQTPSPFEFTTELDGVRHLHFAPCKPLLRBEVSFRVGLHEYFVGSQLPCEPE 2165
Qy      1141 PDVAVLTSMLTDPGSHITAETAKRELARGSPSLASSASQLSAPLSKATCTTHRSDPAD 1200
Db      2166 PDVAVLTSMLTDPGSHITAETAKRELARGSPSLASSASQLSAPLSKATCTTHRSDPAD 2225
Qy      1201 LIEANLWRQEMGNGNITRVESENKVVILDSFEPLQAEEDEREVSVPAAEILRSRKFPRAM 1260
Db      2226 LIEANLWRQEMGNGNITRVESENKVVILDSFDPPLVAEEDEREVSVPAAEILRSRFPAPAL 2285
Qy      1261 PIWARPDPNPLLESWKDPDYPVHVHGCPLPPAKAPPIPPPRKRRTVVLSESTVSSALA 1320
Db      2286 PWARPDYNPLLVETWKKPDYEPVHVHGCPLPPPRSPVPPPRKKRTVWLTESTLPTALA 2345
Qy      1321 ELATKTFGSSSAVDGSTATASDPQSDDG-DAGSDVESYSSMPPLEGPDPLSDGS 1379
Db      2346 ELATKFGSSSTSGI-TGDNITTSSEFAPSGCPDSDVESYSSMPPLEGPDPLSDGS 2404
Qy      1380 WSTVSEBA-SEDVVCCSMSYTWGTALITPCAABETKLPINALSNSLRHNLVYATTSRS 1438
Db      2405 WSTVSSGADTEDVVCCSMSYTWGTALVTPCAABEQKLPINALSNSLRHNLVYATTSRS 2464
Qy      1439 ASLRQKVTRDRLOVLDHDDVLEKMAKASTVKAKLLSVEBEACKLTPPHSARSKFGYG 1498
Db      2465 ACQRKKKVTDRLOVLDHDDVLEKMAKASTVKAKLLSVEBEACSLAPHSARSKFGYG 2524
Qy      1499 AKDVRLSSKAVNHRISVWKKLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 1558
Db      2525 AKDVRCARAKAVAHINSVWKKLLEDSVTPIDTTIMAKNEVFCVQPEKGRKPARLIVFPD 2584
Qy      1559 LGVRVCEKMAKYDVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFYDTRCF 1618
Db      2585 LGVRVCEKMAKYDVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKKTPMGLSYDTRCF 2644
Qy      1619 DSTVTENDIRVEESYQCCDLAPARQAIRSLTERLYIGGPLTNSKQNCYRRCRASGV 1678
Db      2645 DSTVTESDIRTEEAIIYQCCDLDPQARVAIKSLTERLYVGGPLTNSRGENCYRRCRASRV 2704
Qy      1679 LTTSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICBSAGTQDEASLRAFTAMT 1738
Db      2705 LTTSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICBSAGVQDEASLRAFTAMT 2764
Qy      1739 RYSAPGDPDPPEYDLELITSCSSNSVAHDGSKRVYLYLTDPTTTLARAAMETARHTP 1798
Db      2765 RYSAPGDPDPPEYDLELITSCSSNSVAHDGSKRVYLYLTDPTTTLARAAMETARHTP 2824
Qy      1799 VNSWLGNIIMVAPTLWARMILMTHFFSILLAOEQLKALDQCIYAGCYSBPLDLPQII 1858
Db      2825 VNSWLGNIIMVAPTLWARMILMTHFFSILLAOEQLKALDQCIYAGCYSBPLDLPQII 2884
Qy      1859 RLHGLSAFSLHSYSGEINRVASCLRLKLGVPPLRVWRHRSVRARLLSQSGRAATCGKY 1918
Db      2885 RLHGLSAFSLHSYSGEINRVAAACRLKLGVPPLRAWRHRSVRARLLSQSGRAATCGKY 2944
Qy      1919 LFNWAVRTKLTPIPAASQDLSSWFFVAGYSGGDIYHLSRARPRFWMCLLSLVGVG 1978
Db      2945 LFNWAVRTKLTPIPAASQDLSSWFFVAGYSGGDIYHLSRARPRFWMCLLSLVGVG 3004
Qy      1979 IYLLPNR 1985
Db      3005 IYLLPNR 3011
```

RESULT 9

```

JC5620
genome polyprotein - hepatitis C virus (isolate EUH1480)
N;Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C;Species: hepatitis C virus
C;Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C;Accession: JC5620
N;Chamberlain, R.W.; Adams, N.J.; Taylor, L.A.; Simmonds, P.; Elliott, R.M.
Biochem. Biophys. Res. Commun. 236, 44-49, 1997
A;Title: The complete coding sequence of hepatitis C virus genotype 5a, the predominant
A;Reference number: JC5620; MUID:97366593; PMID:9223423
A;Accession: JC5620
A;Molecule type: mRNA
A;Residues: 1-3014 <CHA>
A;Cross-references: UNIPROT:O39928; GB:Y13184
A;Experimental source: genotype 5a, which predominates in South Africa
A;Note: the translation of the nucleotide sequence is not complete in this paper
C;Superfamily: hepatitis C virus genome polyprotein
C;Keywords: ATP; glycoprotein; hydrolase; nucleotide binding; P-loop; polyprotein; serine
F;2-115/Product: capsid protein C #status predicted <CPC>
F;116-191/Product: envelope protein M #status predicted <BPM>
F;192-389/Product: major envelope protein E #status predicted <MEE>
F;384-408/Region: hypervariable #status predicted
F;390-730/Product: nonstructural protein NS1 #status predicted <NS1>
F;731-1007/Product: nonstructural protein NS2 #status predicted <NS2>
F;1008-1616/Product: hepatitis C virus nonstructural protein NS3 #status predicted <NS3>
F;1231-1238/Region: nucleotide-binding motif A (P-loop)
F;1313-1318/Region: nucleotide-binding motif B
F;1317-1320/Region: DEXH motif
F;1617-1863/Product: nonstructural protein NS4a #status predicted <N4A>
F;1864-2014/Product: nonstructural protein NS4b #status predicted <N4B>
F;2015-3014/Product: nonstructural protein NS5 #status predicted <NS5>
F;2210-2245/Region: interferon sensitivity determining #status predicted
Query Match 80.4%; Score 8412.5; DB 1; Length 3014;
Best Local Similarity 77.6%; Pred. No. 0;
Matches 1545; Conservative 202; Mismatches 236; Indels 7; Gaps 3;
Qy      1 MAPITAYSQTRGLGCIITSLGRDRNQVEGEVQVSTATQSFATCVNGVCMTVYHGA 60
Db      1027 LAPITAAQQTGRVGLGAILSLTGRDKNEAGEVQLSTATQTFLGICINGVMWTLFHA 1086
Qy      61 GSKTLAGPKPIQMTYNNVDOLVGWQAPGASLTCTCGSSDLYLVTRHADVIPRRR 120
Db      1087 GSKTLAGPKPVQMTYNNVDKDLVGWSPGPKSLTCTCGSADLYLVTRHADVIPARR 1146
Qy      121 GDSRGSLLSPRVSYLKSGSGGPLLCPSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180
Db      1147 GDTFASLLSPRPISYLKGSSGGPIMCPSHGVGVFRAAVCTRGVAKALEFVVENLETTM 1206
Qy      181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSOKSTKVPAAYAAQGYKVLVLPNSVAATLFGF 240
Db      1207 RSPVFTDNTSPPAVPFHFQVGHLPHTGSGKSTKVPAAYAAQGYKVLVLPNSVAATFGF 1266
Qy      241 AYMSKAGIDPNRTGTGRTTITTCAPTITGYGKFLADGCGSGGAYDIIICDECHSTSTT 300
Db      1267 AYMSRAGVDPNRTGTGRTTITTCAGITITGYGKFLADGCGSGGAYDIIICDECHSQDATT 1326
Qy      301 ILGIGTVLDDQAEATAGARLVVLATATPPGSVTVPHNPNEEVALSGTGPIFYGKAIPETI 360
Db      1327 ILGIGTVLDDQAEATAGARLVVLATATPPGSVTVTHPNIEEVALPSEGIPIFYGRAIPLVI 1386
Qy      361 KGRGHLIFCHSKKKDELAALKSLGLNVAAYYRGLDVSVIPTSGDVIVVATDALMTGFT 420
Db      1387 KGRGHLIFANQKKKAKETAKMKMKPKGRKAVAYYRGLDVAVIPATGDDVVCSTDALMTGFT 1446
Qy      421 GDPSVDICNTCTVOTQTFSLDPTFTTETTTTTPQDAVSRQRRCRTGRGMYRFTFG 480
Db      1447 GDFPSVIDCNSAVTQTQTFSLDPTFTTETTTTTPQDAVSRQRRCRTGRGSHGIRYVSSG 1506
Qy      481 ERPSGMFDSVLCBCYDAGCAWYELTFAETSRLRAYLNTPGLPVCDHLEFWESEVFTGL 540
Db      1507 ERPSGIFDSVLCBCYDAGCAWYELTFAETSRLRAYLNTPGLPVCDHLEFWESEVFTGL 1566
```

```

541  THIDAHFLSOTKAGDNFFVLVAYQATVCARAQAPPPSWDMWKCLIRLKPITLHGTPLL 600
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1567  THIDAHLSQAKGGNFYLVAYQATVCRAKAPPPSWDMTWKMCILKPTLUTGTPLL 1626
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
601  YRLGAVQNEVTTTHPTTKIMACMSADLEVTSTWLVGVGVAALAAAYCLITGTVIVGR 660
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1627  YRLGAVQNEITLTHPTTKIMACMSADLEVTSTWLVGVGVAALAAAYCLITGTVIVGR 1686
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
661  IILSGKPAIIPREVLVYRPDEMEBCASHLYPYEQMGQIAEQPKQAIGLLOTFATQABA 720
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1687  IILSGRPATIPREVLVYQOQDEMEBCASLPYVDEARATAGQPKKVLGLIGTAGQKAE 1746
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
721  AAPVVSXKWTLEAFWAKHMWNTISQIYLAGLSTLPGNPATASLMAFTASTISPLTQ 780
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1747  LKPAATSMXKAEQFWAKHMWNEVSGIYLAGLSTLPGNPATASLMAFTASTISPLTQ 1806
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
781  TLLFNILGGVAAQLAPPASAASAFVAGAGTAAAGVSGIGLKVLDILAGYGAGVAGALVA 840
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1807  TLLFNILGGVWASQIAAPPTAATAFVWVGMAAGVGNIGLGRVLDILAGYGTGVAGALVA 1866
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
841  FKVMSEMESTEDLVNLLFAILSPGALVGVVCAAILRRHVGGEGAVQWNRLLIAFASR 900
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1867  FKIMCGERPTAEELVNLPSILCPGALVGVVCAAVLRRHIGRGEGAVQWNRLLIAFASR 1926
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
901  GNHVSPTHVVPESDAARVTOILSLTITOLKRLHOWINEDCSTPCSGSWLRDVMWLC 960
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1927  GHGSHPTHVVPESDAARVTOILSLTITOLKRLHOWINEDCSTPCSGSWLRDVMWLC 1986
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
961  TVLTDFKTLQSKLLPRLPGVPFFCQRYKGYWRGDGIMQTTCPGAGITGHVKNKSMR 1020
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1987  TALTDKAMLAQKLLPQLPGVPFFCQRYKGYWRGDGIMQTTCPGAGITGHVKNKSMR 2046
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1021  IVGPRTCSNTHGTFFINAYTTCPTSPAPNYSRALRVAABEYVEVTRVGDFFHVTGM 1080
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2047  IVGPKLCSNTHGTFFINAYTTCPTSPAPNYSRALRVAABEYVEVTRVGDFFHVTGM 2106
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1081  TTDNVKCPQVPAPEFTEVDGVRHLRYAPACKFLLREVTFLVGLNOLVNGSQLPCPE 1140
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2107  TQDNLCPCQVSPFEFTLGDVIRHFPAPCNPLREVTFSVGLHVSVVGSQLPCPE 2166
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1141  PDVAVLTSLMTPSHITAEAKRLARGSPPSLASSASQLSAPLSKATCTTRHDSFAD 1200
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2167  PDVAVLTSLMTPSHITAEAKRLARGSPPSLASSASQLSAPLSKATCTTRHDSFAD 2226
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1201  LIEANLLWQMGNTIRVESENKVVILDSFEPLOABEDEREVSVPABILRRSRKPPRAM 1260
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2227  LIEANLLWQMGNTIRVESENKVVILDSFEPLOABEDEREVSVPABILRRSRKPPRAM 2286
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1261  PIWARPDPNPPLESWKDPDYVPVVGCPPLPAKAPPIPPPRKR-TVVLSESTVSSAL 1319
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2287  PWAREGYDPPLELTWKRDYDPPQVWGCPPIPPAGPPVPLPRKPKPMELSDSTVSQM 2346
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1320  AELATKTF---GSSSSAVDSGTATASDPQSDGSDGSDVSEYSSMPPEGEPPDPL 1375
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2347  ADLADARFVDTFPISEIGQDSALGTSQSDSGEPEKRDNDSDAASVSMPPPEGEPPD 2406
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1376  SDGWSVTGSEASEDVVVCCSMSTWTGALITPCAAEETKLIPINALSNLLRHNLVYAT 1435
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2407  SSGSWSTVSGE--DNVCCSMSTWTGALITPCAAEETKLIPINALSNLLRHNLVYAT 2464
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1436  SRSASIRQKVTDFRLQVLDDHRYDLKEMKAKASTVKAKLJSVEACKLTPPHSARSKF 1495
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2465  SRSAGLRQKVTDFRLQVLDDHRYDLKEMKAKASTVKAKLJSVEACKLTPPHSARSKF 2524
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1496  GYGAKDVRLSSKAVNHIKSVWDLLEDTETPTDITIMAKNEVFCVQPKGKPKPARLIV 1555
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2525  GYGAKVRLSKKALHIEGVWQDLDDSDTLEPTTITIMAKNEVFAVEPSKGGKPARLIV 2584
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1556  FPDGLVRVCEKMALYDVVSTLPQAVMGSSYGFQVSPQORVETLVNAWKAKCPMGAYDT 1615
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2585  YPDGLVRVCEKMALYDVVSTLPQAVMGSSYGFQVSPQORVETLVNAWKAKCPMGAYDT 2644
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1030  LAPITAYAOQTGLLGTIVVSMTRDKTEQAGIEQLVLTSTVTSFGLTTSIGVLTWYHGA 1089
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61  GSKTLGAPKGPITOMVTNVDDQLVGMQAPPGARSITPCCTCGSDLVLTNRHADVIPRR 120
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1090  GNKTLGASRGVPTQMTYSSAEGDLVGMPSPPGKTSLEPCTCGAVDLVLTNRHADVIPARR 1149
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1616  RCFDSTVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNGCYRRCA 1675
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2645  RCFDSTVTENDIRVESIYQCCDLAPEARQAIRSLTERLYIGGLTNSKQNGCYRRCA 2704
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1676  SGVLTTSCGNTLTCYLKAAAACRAAKLODCTMLVCGDDLVVICESAGTQDEASLAFTE 1735
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2705  SGVFTTSMGNTWTCYIKALASCAAKLRDCTLLVCGDDLVVICESAGTQDEASLAFTE 2764
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1736  AMTRYSAPGDDPKPYDLELTITSCSSNVSVAHADAGKRVVYLTROPTTPPLAANAETAR 1795
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2765  AMTRYSAPGDDPKPYDLELTITSCSSNVSVAHADAGKRVVYLTROPTTPPLAANAETAR 2824
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1796  HTPVNSWLNINIMVAPTLWARMILMTHFFSIIIAQOLEKALDCQIYGCYSTEPDLQ 1855
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2825  HTPVNSWLNINIMVAPTLWARMILMTHFFSIIIAQOLEKALDCQIYGCYSTEPDLQ 2884
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1856  IITORLHGLSAFSLHYSYSGEINRVSACLRKGLVPPPLRVWRHARSVRARLLSOGGAAATC 1915
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2885  IITORLHGLSAFSLHYSYSGEINRVSACLRKGLVPPPLRVWRHARSVRARLLSOGGAAATC 2944
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1916  GKYLENWAYVTKLKTPIPAASQDLSSWVAGYSGDLYHLSRARPFRMFWKCLLLSV 1975
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2945  GKYLENWAYVTKLKTPIPAASQDLSSWVAGYSGDLYHLSRARPFRMFWKCLLLSV 3004
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1976  GVGIYLLPNR 1985
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3005  GVGIFLLPAR 3014
QY      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
RESULT 10
QJ1303
genome polyprotein - hepatitis C virus (isolate HC-J6)
N:Contains: capsid protein C; envelope protein M; hepatitis C virus (nonstructural) protein NS4a; nonstructural protein NS4b; nonstructural protein NS5
C:Species: hepatitis C virus
C:Date: 19-May-2000 #sequence_revision 19-May-2000 #text_change 09-Jul-2004
C:Accession: JQ1303
R:Okamoto, H.; Okada, S.; Sugiyama, Y.; Kurai, K.; Iizuka, H.; Machida, A.; Miyakawa, Y.
J. Gen. Virol. 72, 2697-2704, 1991
A:Title: Nucleotide sequence of the genomic RNA of hepatitis C virus isolated from a human
A:Reference number: JQ1303; MUID:92044440; PMID:1658196
A:Accession: JQ1303
A:Molecule type: Genomic RNA
A:Residues: 1-3033 <OKA>
A:Cross-references: UNIPROT:P26660; GB:D00944; NID:G221650; PID:BAA00792.1; PID:G221651
A:Experimental source: isolate HC-J6 from a Japanese individual
C:Superfamily: hepatitis C virus genome polyprotein
C:Keywords: ATP; glycoprotein; hydrolase; P-loop; polyprotein; serine proteinase; transmembrane
F2-115/Product: capsid protein C #status predicted <CPC>
F2-116-191/Product: envelope protein M #status predicted <EPM>
F192-389/Product: major envelope protein E #status predicted <MEE>
F390-733/Product: nonstructural protein NS1 #status predicted <NS1>
F734-1010/Product: nonstructural protein NS2 #status predicted <NS2>
F1011-1619/Product: hepatitis C virus #status predicted <NS3>
F1316-1321/Region: nucleotide-binding motif B
F1320-1323/Region: DEH motif
F1620-1866/Product: nonstructural protein NS4a #status predicted <NS4a>
F1867-2017/Product: nonstructural protein NS4b #status predicted <NS4b>
F2018-3033/Product: nonstructural protein NS5 #status predicted <NS5>
F196,209,234,305,325,417,423,430,448,477,534,542,558,578,627,649,1091,1217,1259,2038,2828
```

Qy	121	GD	R	G	S	L	L	S	R	P	P	V	S	Y	L	K	G	S	G	P	L	L	C	P	S	G	H	A	V	G	I	F	R	A	A	V	C	T	R	G	V	A	K	A	V	D	F	V	P	V	E	S	M	E	T	T	M	180
Db	1150	G	D	R	G	A	L	L	S	R	P	P	I	S	T	L	K	G	S	G	P	P	L	C	P	R	G	H	A	V	G	V	F	R	A	A	V	C	T	R	G	V	A	K	I	P	I	V	E	T	D	I	V	T	1209			
Qy	181	R	S	P	V	E	T	D	N	S	P	P	A	V	P	O	N	F	O	V	A	H	L	H	A	P	T	G	S	K	T	V	P	A	A	A	O	G	K	V	L	V	L	N	P	S	V	A	A	T	L	G	F	240				
Db	1210	R	S	P	T	S	D	N	S	T	P	P	A	V	P	O	T	Y	Q	Y	L	H	A	P	T	G	S	K	T	V	P	A	A	A	O	G	K	V	L	V	L	N	P	S	V	A	A	T	L	G	F	1269						
Qy	241	A	Y	M	S	K	A	H	I	D	N	I	R	T	G	V	R	T	I	T	G	A	P	I	T	S	T	G	K	L	P	A	I	T	S	T	G	K	L	P	A	I	T	S	T	G	K	L	P	A	I	T	S	T	300			
Db	1270	A	Y	L	S	K	A	H	I	D	N	I	R	T	G	V	R	T	I	T	G	A	P	I	T	S	T	G	K	L	P	A	I	T	S	T	G	K	L	P	A	I	T	S	T	G	K	L	P	A	I	T	S	T	1329			
Qy	301	I	L	G	T	V	L	D	O	A	E	T	A	G	A	R	L	V	L	A	T	A	T	P	G	S	V	T	V	P	H	N	T	E	E	V	A	L	S	T	G	B	E	I	P	Y	G	K	A	I	P	E	T	360				
Db	1330	I	L	G	T	V	L	D	O	A	E	T	A	G	A	R	L	V	L	A	T	A	T	P	G	S	V	T	T	P	H	N	T	E	E	V	A	L	S	T	G	B	E	I	P	Y	G	K	A	I	P	E	T	1389				
Qy	361	K	G	R	H	L	F	C	H	S	K	K	C	D	E	L	A	K	L	S	G	L	H	N	A	V	A	Y	R	G	L	D	S	V	I	P	T	S	G	D	V	I	V	A	T	A	L	M	T	G	F	420						
Db	1390	K	G	R	H	L	F	C	H	S	K	K	C	D	E	L	A	K	L	R	O	N	G	L	N	A	V	A	Y	R	G	L	D	S	V	I	P	T	O	G	D	V	V	V	A	T	A	L	M	T	G	F	1440					
Qy	421	G	D	F	S	V	I	D	C	N	T	V	T	O	D	F	S	L	D	P	T	I	E	T	T	T	V	P	O	D	A	V	S	R	O	R	R	G	R	T	G	R	G	M	G	I	R	F	V	T	P	G	480					
Db	1450	G	D	F	S	V	I	D	C	N	A	V	T	O	D	F	S	L	D	P	T	I	E	T	T	T	T	V	P	O	D	A	V	S	R	O	R	R	G	R	T	G	R	G	I	R	Y	V	Y	S	T	G	1509					
Qy	481	E	R	S	G	M	F	D	S	S	V	C	B	Y	D	A	G	A	N	E	L	T	P	A	E	T	S	R	V	L	R	A	Y	L	N	T	P	G	L	P	V	O	D	H	L	E	F	M	E	S	V	T	G	L	540			
Db	1510	E	R	A	S	G	M	F	D	S	V	V	L	C	B	Y	D	A	G	A	N	E	L	T	P	A	E	T	T	V	L	R	A	Y	F	N	T	P	G	L	P	V	O	D	H	L	E	F	M	E	A	V	T	G	L	1569		
Qy	541	T	H	I	D																																																					

[illegible]







Db 771 IWACMSADLEVTSTWLVGVVAALAAVCLTVGSVAIVGRILSGKPAIIDREALYQ 830  
 QY 680 FMEBECASHLYIYQGMOLARQFKQKALGLLOATKQABAAAP 723  
 Db 831 FMEBECASLPYMDTRAIAQGFKEKVLGFIITGQKAEITLK 874  
 RESULT 14  
 T08841  
 polyprotein - douroucouli hepatitis GB virus A  
 C:Species: douroucouli hepatitis GB virus A  
 C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 17-Nov-2000  
 C:Accession: T08841  
 R:Erker, J.C.; Desai, S.M.; Leary, T.P.; Chalmers, M.L.; Montes, C.C.; Mushahwar, I.K.  
 J. Gen. Virol. 79, 41-45, 1998  
 A:Title: Genomic analysis of two GB virus A variants isolated from captive monkeys.  
 A:Reference number: Z16486; MUID:98120818; PMID:9460920  
 A:Accession: T08841  
 A>Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-3005 <ERK>  
 A:Cross-references: EMBL:AF023425; NID:g2828599; PIDN:AAC40502.1; PID:g2828600  
 C:Superfamily: hepatitis C virus genome polyprotein  
 C:Keywords: polyprotein  
 Query Match 19.9%; Score 2078.5; DB 2; Length 3005;  
 Best Local Similarity 31.2%; Pred. No. 4.9e-118;  
 Matches 647; Conservative 288; Mismatches 868; Indels 271; Gaps 70;  
 QY 2 APITAYSQOTRGLGCIITSLTRDRNOVEGEVQVNSTATQFLACVNGVCMVTVHGAG 61  
 Db 979 APVVV-MQRLGFFSVKTSMLGRDERHEGSIVLGTSTTRSMGTGVNGVMYTFHGSN 1037  
 QY 62 SKTLAPKGPITOMYNTVDQVLVGMQAPPGARSLPCTCGSDLYLVTRHADVIPVRRKG 121  
 Db 1038 ARTLAGPVGVNCRWMSRDDVAVPLPGSCASCLEPCKGTQSQVWCIRN--DGLCHGRL 1095  
 QY 122 DSRGILLSRPVSYLKGSGGLPCPSGHAVGIFRAAVCTRGVAKAVDFVPVSEMTWR 181  
 Db 1096 SKLVELDLTEISDRFGSGSLPCDEGHVGMN--VSLHVRGV-KTVGVRYVKPWETLEPK 1153  
 QY 182 SPVFTDNSSFPAPVQT-FQVAHLHAPTGGSKTKVPAAYAAQYKVLVNLPSVAATLFGF 240  
 Db 1154 DSQ-VKSEAPPVPGTKGTFTAPLYLFTGSGKSTRIEYTKAGHKVLVNLPSIATVRAG 1212  
 QY 241 AYMSKAHGDPIRTGVRII---TTGAPITTSYTKFLADGGSCGAVDIIICDCHST 296  
 Db 1213 PYMEKLSGQHSIYCGHDTTAYSRITGSLTYCTYGRFMANPRRYLRGADIVICDCHVT 1272  
 QY 297 DSTTILGIGTVLDQAGTAGARLVVLATATPPGSVTVPHNIEVALSSTGEIPFYGKAIP 356  
 Db 1273 DPTSLVGMGRALLARECGVRLILFATATPPGAPLAQHSIREVPLGVGDGEVAFYGHKUP 1332  
 QY 357 IETIKGRHLIFCHSKKCDLAAKLGLGLNNAVAYRGLDVSIVPTSGDVIIVATDALM 416  
 Db 1333 VERYRTGRHLRCHSKVCECNRLHAALS TAGCNAVYVRGNEQEI--PAGDVCVCADALS 1390  
 QY 417 TGTGDPDSVIDNCTVTVDFSLDPTTETIETTPQDAVRSQRGRGTGRGMGIYRF 476  
 Db 1391 TGYTGGSFTVTCGLMVESEVETLDTITISVRTTPAPAEALRAQRGRGSGQGTYYY 1450  
 QY 477 VTPGERPSGMFSSVLCECYDAGCAWVLTPTAETS VRLRAYLNTPGLPVCQDHLPEWESV 536  
 Db 1451 AMTASAPAGTLRSGPLWAAVEAGVAVNLEPDMWTADILLRAYDACPTTAITASVGEAINF 1510  
 QY 537 FTGLTHIDAH-FLSQTKAGDNFPYLVAYQATVCARQAAPPSPSWDMQWKLRLKPTLFG 595  
 Db 1511 FSLGLVPMRNPYQVAMAKSHQHNWPLLVGVQVTCMQEANAVAGPGDGPWAGLACT-----G 1565  
 QY 596 PTLILYRLGAVQNEVTTTHPIKYMA-----CMSADLEVTVSTWLVGVGLAAL 645  
 Db 1566 PIFLLCRWGARPSPVAPHHVVDDLOARLGVABEGYSPCYAGPI-----LLVGLALAGG 1618

QY 646 AAYCLTTGTSVIV--GRILSGKPAI-----IPDREVLVREFDEME---ECASHL 690  
 Db 1619 AVLAWHTGSLVVVTSWRVNGNGNPLIQOSTRQVSTSAFYQLTVCEGEQTPADGKCAAA 1678  
 QY 691 PYIEQGM---QLAEQFKQKAI-GLLOATKQAEAAAPVVEKSWRTL-----732  
 Db 1679 VOLXESTCGWGPMAASFDCAGMKGVLDNSMRTTAAAAAVEKSDSLMRSFCANNYCPGGGAT 1738  
 QY 733 --EAFWAK-----HWN--FISGTOYLAGLSTLPG---NPAIASLMAFTASITSLPTTQ 779  
 Db 1739 SASAFASLDTKFAQAMDIFTNGRSLVLVGVAGYARRNPPLGVAAAFMGXSGAGHVH 1798  
 QY 780 HTLLENILGGMVAALAPPSAASAFVAGIAGAAVGSIGLKVLDVILLAGYGAGVAGALV 839  
 Db 1799 VFLAALLLGVCGTMLGTPSVGLMSGAYFAGSITSSWL-SATVAVLGGWEGAXNAASL 1857  
 QY 840 AFKVMGSEMPSTEDLVNLLPAILSPGALVGVVCAAIL-RRHVGPGEVAVQVNNRLIAFA 898  
 Db 1858 TFDFLTGR-AELKDMWFLVSCXASPGASVAGVALGILLMSMKKGVE---DWNRLITLL 1913  
 QY 899 SRGNHVSPTHYVPESDAARVITQILSSLITOLLKELHOMINEDCSTPCSGSWLRDVWVW 958  
 Db 1914 PRGS-VLPDGFVKSEFTERVSTILRKMSLSRWMTLVERRELDLLETPCS-SMLMDLIDW 1971  
 QY 959 ICTVLTDFKTLQSKLLPRLPG--VPFFSCQRYGVWRGDMOTTCPCGAQITGHVKN 1016  
 Db 1972 ----LVRGRIYGRRLKGMVPSVRVPLVCGTPEGWGSVGEHIEARCAACGCIITADVEE 2027  
 QY 1017 GSMRIV--GPRTCSTNHTGTFPINAYTTGCPPTS-PAPNYSRALMRVAABEVETVRGD 1073  
 Db 2028 GKLDVHYSSRLCSNVLKGTVPVSAAGSDAEPEVPA---GPALYIGVAEWQLVR-KD 2083  
 QY 1074 FHYVTGMT-----DNVKCPQOVAPPEFTVDGLVRLHRYAPACKPLRBEETFLVGLN 1127  
 Db 2084 KTLVVGASSVHLHIDELRRAIRGP-PMF---VGVGVSWEAFLQOQPLP-----2128  
 QY 1128 QYLVGSQ-----LPCEPEPDVAVLT-----SMLTDPHSI 1156  
 Db 2129 VYRAGSVRFDDVYRSLPHITLALPPPPPPPLAMPPPPPPEQVWTEBEEADLRBAR 2188  
 QY 1157 TAETAKRLIARGSPSL-ASSASQLSAPSL-----KATCTTR--HDSPPADLI 1202  
 Db 2189 AIEAVNERLPVNPPEAAQALDALBEAAVSLPHVGAIIIGDCSCESESGFHIFPEPDAV 2248  
 QY 1203 EANTLWRQEMGNI---TEVESENKVILDSPELQAEEDEREVSVP---AEILRSRK 1255  
 Db 2249 DVFIGMVEVQVGLRDQARDLGLRLAVLGLARLESABAEHPEASLNTERTMGEIDTLAD 2308  
 QY 1256 FPRAMPIMARPDPNYPPLLESWMKDPDVPVPV--HGCLPLPPAKAPPI-----PPRRKR 1306  
 Db 2309 VOARLDVTCRSDTSGSSFEQISLSDSEPETIVEGGLKLEVRPQPVRFKDLIRPGEGAKR 2368  
 QY 1307 TVVLESTVSSALAEATKTPGSSSSAVDSGTATPAPQDDGAGDVSYSNMPPL 1366  
 Db 2369 LVTVRQSCCAD--RSATRAF--ALSLPIAAVATILSFOL-TDHTVSDTGRVLDPLELL 2422  
 QY 1367 EGEPGDPDLSGWSWSTVSEASEDVVCCSMYTWTCALI-----TPCAAF 1411  
 Db 2423 QNAVGD-----LIVACRRSESV--SYSIWSGAPLNTGRHOPAPMTRPIGHTITAD 2472  
 QY 1412 ETKLPIINALSNLRLHNLVYATTSRSLRQKVTFFDR-LOVLDDHYDVLKEMKAKAS 1470  
 Db 2473 TTK-----VYVTDNQAQERAAKVTIRGSRVYDAHYGVVSEVLQQAQ 2516  
 QY 1471 TVYAKILLSVEACKLTPPHSARSKFYGAK-DVRNLSKA-----VNHRSVKKLLLEDTE 1525  
 Db 2517 TIKSPGWYDEA--IAKVHS-RAAAGFGSKVTVGNMTTPAARAEVNNMLAKIK-----TK 2568  
 QY 1526 TPDITTIMAKNVCVQPEKGRKPARLIVFDPLGVRCVKMALYDVVSTPLPQAVMGSSY 1585  
 Db 2569 QEVPTFLVTKREVPF---OKTTRKPRFICFPPLDFRIAEKMLIGD-PGLVAKGILGKSY 2624



QY 1331 -----ESSAVDSGTATASPDQDDG-----AGSDVESYSSMPPLLEGEPDPLSDGSM 1380  
 Db 2337 RILRVQGCCDRSTLIKAFPLE-LKVGELFVCGYDPEGHR-----VTDDNGDDMN----- 2387  
 QY 1381 STVSBEASEDDVCC-----SMSTWTGA-----LITPCAAEBTKLPINALSNSLLR 1426  
 Db 2388 TFIADTLGDIILSCETREBAQXSYSYWSGAPLGCGRVVP-----PITRPIGTHLT 2439  
 QY 1427 HHNL-VYATTSASLRQKVTDFDLQ-VLDHYRDVLKEMKAKASTVAK-----LLSVE 1480  
 Db 2440 HDTTKVYVTDPPRAABRAEKVTIWRQRYDEHYASVWGEALKKAATTPGWTYMAIS 2499  
 QY 1481 EACKLTPPHSARKFGYGAQOVNLSKAVNHRSVWKDLLEDTEPIDTTIMAKNEVFC 1540  
 Db 2500 KYRK-----RAAGYGSKVTVATLETGWPHVRIMDKIRCHEEVPF--TFVTKREVPF 2550  
 QY 1541 VOPEKGRKPARLIVPPDLGVRVCKRMALYDVVSTLPQAVMGSSYGFQYSPGQVFEFLVN 1600  
 Db 2551 ---SKTRKPPRFIVVPDLDFIAEKMILGD-PGIVAKAILGPAYLFQVTPNQRVKLLVS 2606  
 QY 1601 AWKAKCPMGFAVDRCFSTVTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPL 1660  
 Db 2607 TWESKTHPAITVDATCFDSDIDERDMEVEAEIFAASDQPELVKALC----RYAEGPM 2662  
 QY 1661 TNSKGQCGYRRCRASGVLTTCGNLTCTYLKAAAAACRAAKLODCTMLVCGDDLVVICES 1720  
 Db 2663 VTPDGVPIGFRKCRSGVLTSSANSITCYIKVKAACAKVGLKDPSPFFIAGDDCLIIYED 2722  
 QY 1721 AGTQDEASLRAFTEAMTRYAPPDPKPEYDLELITS-CSSNVSVAHDAKRVYIYL 1779  
 Db 2723 DGADPCE-RURL---ALGNY---GYRCPKSHASLDTAECCSAYLAECTAGGSRRWLS 2774  
 QY 1780 RPTTFLARAAMETARHTPVNSWLGNIIMVAPTLWARMILMTHFFSILLAQE---OLEKA 1836  
 Db 2775 TDMRKELARAAAEYS--DFVGSALGTILMYPWHPVIVRYVLIPIHV--LIMAFRGGGTPDEL 2830  
 QY 1837 LDCQIYGACYSIEPLDLPQIQLHGLSAPSLHSYSPGEINRVASCLRLGVPLRVWRH 1896  
 Db 2831 VMQVQGNYSFPLKILPVLVSLHGPWCLQVTTDSTKTMEAGNALRDLGKXLSWHR 2890  
 QY 1897 RARSVRARLLSQG---GRAATCGKYLEFNWAVTKLKLTPIPAASQLDLSWVYAGSGD 1953  
 Db 2891 RAGNVTRILLRGKAWGHLARC---LLWHPGLKEHPPPVIALPNFRL----VTPVEHHE 2942  
 QY 1954 IYHLSRAPRPRWMC 1970  
 Db 2943 EVLISLKERPPWVKWLL 2959

Search completed: December 8, 2004, 12:33:12  
 Job time : 70 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: December 8, 2004, 12:14:57 : Search time 272 Seconds  
(without alignments)  
4198.966 Million cell updates/sec

Title: US-09-576-989-3

Perfect score: 10465

Sequence: 1 MAPITAYSQTRGLGCIIT.....FMWCLLLLSVGVGYLLPNR 1985

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt\_02.\*

1: uniprot\_sprot.\*

2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10462	100.0	3010	2	Q9WMX2
2	10365	99.0	3010	2	Q9QP06
3	10245	97.9	3010	2	Q9J3H2
4	10225	97.7	3008	2	Q9J3F4
5	10225	97.7	3010	2	Q8V638
6	10218	97.6	3010	2	Q93016
7	10218	97.6	3010	2	Q9DTE7
8	10206	97.5	1984	2	Q7T4V8
9	10199	97.5	3010	2	P90194
10	10194	97.4	3010	2	Q99AU2
11	10187	97.3	3010	2	Q9DTE9
12	10183	97.3	3010	2	Q9QIY8
13	10182	97.3	3010	2	P90193
14	10177	97.2	3010	2	Q9J3H3
15	10177	97.2	3010	2	Q9QIX6
16	10177	97.2	3010	2	Q9DIF0
17	10176	97.2	3010	2	Q9J3G7
18	10176	97.2	3012	2	Q9WIK7
19	10172	97.2	3010	2	Q9QIX5
20	10172	97.2	3010	2	Q9QIY4
21	10172	97.2	3010	2	Q9DTE6
22	10170	97.2	3010	1	FOLG_HCVBK
23	10169	97.2	3010	2	Q9QIY3
24	10165	97.1	3010	2	Q9QIY7
25	10165	97.1	3010	2	Q9DTE6
26	10161	97.1	3010	2	Q9J3H9
27	10160	97.1	3010	2	Q9DTE2
28	10159	97.1	3010	1	FOLG_HCVJT
29	10157	97.1	3010	2	Q91AU0
30	10151	97.0	3013	2	Q9QIY0
31	10151	97.0	3014	2	Q6GYR8

32	10150	97.0	3010	2	Q9WIK8	Q9wik8 hepatitis c
33	10149	97.0	3010	2	Q9J311	Q9j311 hepatitis c
34	10148	97.0	3010	2	Q6GYR9	Q6gyr9 hepatitis c
35	10147	97.0	3010	2	Q8Q7P3	Q8q7p3 hepatitis c
36	10147	97.0	3010	2	Q9J3G6	Q9j3g6 hepatitis c
37	10147	97.0	3010	2	Q9J310	Q9j310 hepatitis c
38	10145	96.9	3010	2	Q9J3G9	Q9j3g9 hepatitis c
39	10144	96.9	3013	2	Q9QIX9	Q9qix9 hepatitis c
40	10141	96.9	3010	2	Q9J3H7	Q9j3h7 hepatitis c
41	10141	96.9	3010	2	Q9QIY1	Q9qiyl hepatitis c
42	10136	96.9	3010	2	Q9DTE4	Q9dte4 hepatitis c
43	10132	96.8	3010	2	Q9QP61	Q9qp61 hepatitis c
44	10131	96.8	3010	2	Q9QIX2	Q9qix2 hepatitis c
45	10130	96.8	3010	2	O09796	O09796 hepatitis c

## ALIGNMENTS

RESULT 1  
Q9WMX2  
ID Q9WMX2; PRELIMINARY; PRT; 3010 AA.  
AC Q9WMX2;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus type 1b.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus; Hepatitis C virus type 1.  
OX NCBI\_TaxID=31647;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RB;  
RX MEDLINE=99322193; PubMed=10390360;  
RA Lohmann V, Koerner F., Koch J.O., Herian U., Theilmann L.,  
RA Bartschlag R.;  
RT "Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell line.";  
RL Science 285:110-113(1999).  
DR EMBL; AJ238799; CAB46677.1; -  
DR PIR; A61196; A61196.  
DR PIR; P00246; P00246.  
DR PIR; PS0329; PS0329.  
DR HSP; P26663; IUXP.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR GO; GO:0019079; P:viral genome replication; IEA.  
DR GO; GO:0019087; P:viral transformation; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001410; DEAD\_  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV\_core.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR InterPro; IPR000745; HCV\_NS4a.  
DR InterPro; IPR001490; HCV\_NS4b.  
DR InterPro; IPR002868; HCV\_NS5a.  
DR InterPro; IPR002166; HCV\_RdRp.  
DR InterPro; IPR001650; Helicase C.  
DR InterPro; IPR004109; Peptidase S29.  
DR InterPro; IPR009003; Pept Ser Cys.  
DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.





Db 2706 TSCGNTLTCLYLKAAACRAAKLQDCTMLVCGDDLVLVVCESAGTQDEASRAFTTEAMTRY 2765

Qy 1741 SAPGDDPKPEYDLELITSCSNVSVADHASKRYYVYTRDPTTPPLAANAETARHTPVN 1800

Db 2766 SAPGDDPKPEYDLELITSCSNVSVADHASKRYYVYTRDPTTPPLAANAETARHTPVN 2825

Qy 1801 SWLGNIIWYAPTLWARMILMTHFFSILLAOEALDQCIYGCYSIEPLDLPQIQR 1860

Db 2826 SWLGNIIWYAPTLWARMILMTHFFSILLAOEALDQCIYGCYSIEPLDLPQIQR 2885

Qy 1861 HGLSAFSLHSPGPEINRVASCLRLKGLVPPPLRVWHRARSVRARLLSQGGRAATCGKYL 1920

Db 2886 HGLSAFSLHSPGPEINRVASCLRLKGLVPPPLRVWHRARSVRARLLSQGGRAATCGKYL 2945

Qy 1921 NNAVRTKLTPIPAASQDLSSFFVAGYSGDDIYHLSRARPFRWFMWCLLLSVGVGIY 1980

Db 2946 NNAVRTKLTPIPAASQDLSSFFVAGYSGDDIYHLSRARPFRWFMWCLLLSVGVGIY 3005

Qy 1981 LLPNR 1985

Db 3006 LLPNR 3010

RESULT 2

Q9QP06 PRELIMINARY; PRT; 3010 AA.

AC Q9QP06;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Polyprotein.

OS Hepatitis C virus type 1b.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;

OC Hepacivirus; Hepatitis C virus type 1.

OX NCBI\_TaxID=31647;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=RB;

RX MEDLINE=99370154; PubMed=10438800;

RA Koch J.O., Bartenschlager R.;

RT "Modulation of hepatitis C virus NS5A hyperphosphorylation by

RT nonstructural proteins NS3, NS4A, and NS4B.";

RL J. Virol. 73:7138-7146(1999).

DR EMBL; AJ238800; CAB53095.1; -

DR PIR; A61196; A61196.

DR PIR; PQ0246; PQ0246.

DR PIR; PS0329; PS0329.

DR HSP; P27958; IALR.

DR GO; GO:0016021; C: integral to membrane; IEA.

DR GO; GO:0019028; C: viral capsid; IEA.

DR GO; GO:0019031; C: viral envelope; IEA.

DR GO; GO:0005524; F: ATP binding; IEA.

DR GO; GO:0008026; F: ATP-dependent helicase activity; IEA.

DR GO; GO:0003723; F: RNA binding; IEA.

DR GO; GO:0003968; F: RNA-directed RNA polymerase activity; IEA.

DR GO; GO:0008236; F: serine-type peptidase activity; IEA.

DR GO; GO:0005198; F: structural molecule activity; IEA.

DR GO; GO:0006508; P: proteolysis and peptidolysis; IEA.

DR GO; GO:0006350; P: transcription; IEA.

DR GO; GO:0019079; P: viral genome replication; IEA.

DR GO; GO:0019087; P: viral translation; IEA.

DR InterPro; IPR000345; CytC\_heme\_BS.

DR InterPro; IPR001410; DEAD.

DR InterPro; IPR002522; HCV\_capsid.

DR InterPro; IPR002521; HCV\_core.

DR InterPro; IPR002519; HCV\_env.

DR InterPro; IPR002531; HCV\_NS1.

DR InterPro; IPR000745; HCV\_NS4A.

DR InterPro; IPR001490; HCV\_NS4B.

DR InterPro; IPR002868; HCV\_NS5A.

DR InterPro; IPR002166; HCV\_RdRP.

DR InterPro; IPR001650; Helicase\_C.

DR InterPro; IPR004109; Peptidase\_S29.

InterPro; IPR009003; Pept\_Ser\_Cys.

InterPro; IPR002518; Pept\_U39\_HCV\_NS2.

InterPro; IPR007095; RNA\_pol\_DS\_PS.

InterPro; IPR007094; RNA\_pol\_PSVir.

Pfam; PF01543; HCV\_capsid; 1.

Pfam; PF01542; HCV\_core; 1.

Pfam; PF01539; HCV\_env; 1.

Pfam; PF01560; HCV\_NS1; 1.

Pfam; PF01538; HCV\_NS2; 1.

Pfam; PF02907; HCV\_NS3; 1.

Pfam; PF01006; HCV\_NS4a; 1.

Pfam; PF01003; HCV\_NS4b; 1.

Pfam; PF01506; HCV\_NS5a; 1.

Pfam; PF00271; Helicase\_C; 1.

Pfam; PF00998; Viral\_RdRP; 1.

SMART; SM00487; DEXDC; 1.

PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN 1.

Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;

Polyprotein; Transmembrane.

CHAIN 810 1026 NS3 proteinase.

FT CHAIN 1027 1657 NS3/4A proteinase cofactor.

FT CHAIN 1658 1712 core protein.

FT CHAIN 1 191 NS4B protein.

FT CHAIN 1712 1972 NS5A phosphoprotein.

FT CHAIN 1973 2419 NS5B RNA dependant RNA polymerase.

FT CHAIN 2419 3010 glycoprotein E1.

FT CHAIN 192 383 glycoprotein E2.

FT CHAIN 384 746 p7 Peptide.

FT CHAIN 747 809

SQ SEQUENCE 3010 AA; 327000 MW; A570B980DD64634 CRC64;

Query Match 99.0%; Score 10365; DB 2; Length 3010;

Best Local Similarity 98.9%; Pred. No. 0;

Matches 1964; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

Qy 1 MAPITAYSQQTGLGCIITSLTGRDRNOVEGEVQVVTATQSFATCVNGVCMVTHGA 60

Db 1026 LAHITAYSQQTGLGCIITSLTGRDRNOVEGEVQVVTATQSFATCVNGVCMVTHGA 1085

Qy 61 GSKTLAGPKPIITQMTYNVDQDLVGWQAPPQAKSLTPTCTGSSDLYLVTRHADVIPVRR 120

Db 1086 GSKTLAGPKPIITQMTYNVDQDLVGWQAPPQAKSLTPTCTGSSDLYLVTRHADVIPVRR 1145

Qy 121 GDSRGSLLSPRVSYLKSGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180

Db 1146 GDSRGSLLSPRVSYLKSGSGGGLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205

Qy 181 RSPVFTDINSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLSPG 240

Db 1206 RSPVFTDINSPPAVPQTFQVAHLHAPTSGSKSTKVPAAAYAAQGYKVLVLPNSVAATLSPG 1265

Qy 241 AYMSKAHGIDPNIRITGVRTITTCAPTITSTYTGKFLADGGCGGAYDIIICDECHSTDTT 300

Db 1266 AYMSKAHGIDPNIRITGVRTITTCAPTITSTYTGKFLADGGCGGAYDIIICDECHSTDTT 1325

Qy 301 ILGIGTVLDOAETAGARLVVLATATPPGSVTVPHNPNEEVALSSTGEPYGAIPETI 360

Db 1326 ILGIGTVLDOAETAGARLVVLATATPPGSVTVPHNPNEEVALSSTGEPYGAIPETI 1385

Qy 361 KGRHLIFCHSKKKCDLAALKSLGLNNAVAYRGLDVSIVPTSGDIVVATDALTGTFT 420

Db 1386 KGRHLIFCHSKKKCDLAALKSLGLNNAVAYRGLDVSIVPTSGDIVVATDALTGTFT 1445

Qy 421 GFDSDVIDCNTCVTQVDFSLDPTFTIETTTPQDAVSRQRGRGRTGRGMGIYRVFTPG 480

Db 1446 GFDSDVIDCNTCVTQVDFSLDPTFTIETTTPQDAVSRQRGRGRTGRGMGIYRVFTPG 1505

Qy 481 ERPSGMFDSVLCCEYDAGCAMEYELTPAETSVRLRAYLNTPLGIPVQDHLFEWESVFTGL 540

Db 1506 ERPSGMFDSVLCCEYDAGCAMEYELTPAETSVRLRAYLNTPLGIPVQDHLFEWESVFTGL 1565

Qy 541 THIDAHFLSQTQAGDNFFLYVAYQATVTCARAAQAPPPSDQWKKLIRLKPTLHGFTPL 600

1566 THIDAHFISQTKAQGNRPYLVAQATVCARAQAPPPSWDQWQKCLIRLKPTLHGPTPL 1625  
 601 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIVGR 660  
 1626 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIVGR 1685  
 661 IILSKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEOFKQKAIIGLLOATKQAEA 720  
 1686 IIVLSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEOFKQKAIIGLLOATKQAEA 1745  
 721 AAPVUESKWRTEAFWAKEMWFIISGIVLAGLSTLPGNPAIASLMAFTASITSPITTOH 780  
 1746 AAPVUESKWRTEAFWAKEMWFIISGIVLAGLSTLPGNPAIASLMAFTASITSPITTOH 1805  
 781 TLLFNILGWAQAALAPPAAAFVAGAGIAGAAVGSIGLGVLDVILAGYGAGVAGALVA 840  
 1806 TLLFNILGWAQAALAPPAAAFVAGAGIAGAAVGSIGLGVLDVILAGYGAGVAGALVA 1865  
 841 FKVMSEMESTEDLVNLLPAILSPGALVGVVCAALTRRHVGPGEAGVQWMLIAFASR 900  
 1866 FKVMSEMESTEDLVNLLPAILSPGALVGVVCAALTRRHVGPGEAGVQWMLIAFASR 1925  
 901 GNVHSPTHVVPESDAARVTOILSSLTITQLKRLHOWINEDCTPCSGSWLRDVMWIC 960  
 1926 GNVHSPTHVVPESDAARVTOILSSLTITQLKRLHOWINEDCTPCSGSWLRDVMWIC 1985  
 961 TVLTFDKTLQSKLLPRLPGVPPFCQYKGVWRGDMQTTCPGCAQITGHVKNQSMR 1020  
 1986 TVLTFDKTLQSKLLPRLPGVPPFCQYKGVWRGDMQTTCPGCAQITGHVKNQSMR 2045  
 1021 IVGPRTCSNTHGTFFINAYTTCPTSPAPNYSRALNRVAABEYVEVTRVGDVHVTGM 1080  
 2046 IFGPRTCSNTHGTFFINAYTTCPTSPAPNYSRALNRVAABEYVEVTRVGDVHVTGM 2105  
 1081 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLLRREVTFLVGLNQYLVSQLPCEPE 1140  
 2106 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLLRREVTFLVGLNQYLVSQLPCEPE 2165  
 1141 PDVAVITSMITDPSHTAETAKRRLARGSPPSLASSASQLSAPSLKATCTTHDSDPAD 1200  
 2166 PDVAVITSMITDPSHTAETAKRRLARGSPPSLASSASQLSAPSLKATCTTHDSDPAD 2225  
 1201 LIEANILWQEMGNTIRVSEENKVLIDSFEPLQAEEDREVSVPAAEILRRSRKPPRAM 1260  
 2226 LIEANILWQEMGNTIRVSEENKVLIDSFEPLQAEEDREVSVPAAEILRRSRKPPRAM 2285  
 1261 PIWAPDYNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSALA 1320  
 2286 PIWAPDYNPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSESTVSALA 2345  
 1321 ELATKTFGSSSAVDSGTATAPDQPSDDGSDGSDVESYSSMPPLGEGFDGDLSDGSW 1380  
 2346 ELATKTFGSSSAVDSGTATAPDQPSDDGSDGSDVESYSSMPPLGEGFDGDLSDGSW 2405  
 1381 STVSEERASDVCCSMSTYTGALITPCAAEETKLPINALSNLRRHNLVYATTSAS 1440  
 2406 STVSEERASDVCCSMSTYTGALITPCAAEETKLPINALSNLRRHNLVYATTSAS 2465  
 1441 LRQKKVTFDLRLQVLDHVRDVLKEMKAKASTVKAKLSVEEACKLTPPHSARSFGYCAK 1500  
 2466 LRQKKVTFDLRLQVLDHVRDVLKEMKAKASTVKAKLSVEEACKLTPPHSARSFGYCAK 2525  
 1501 DVRLNSKAVNHRSWKDLLEDTEPIDTITIMAKNEVFCVQPEKGGKPARLIVFPDLG 1560  
 2526 DVRLNSKAVNHRSWKDLLEDTEPIDTITIMAKNEVFCVQPEKGGKPARLIVFPDLG 2585  
 1561 VRVCEKMAVYVSTLPOAVMGSSYGFQYSPQGVFLVNAWAKKCPMGFAVYTRCFDS 1620  
 2586 VRVCEKMAVYVSTLPOAVMGSSYGFQYSPQGVFLVNAWAKKCPMGFAVYTRCFDS 2645  
 1621 TVTENDIRVEESIYQCCDLAEPARQAIRSLTERLYIGGPLNLSKQNGCYRRRCASGVLIT 1680  
 2646 TVTENDIRVEESIYQCCDLAEPARQAIRSLTERLYIGGPLNLSKQNGCYRRRCASGVLIT 2705

1681 TSCGNLTCTYLKAAACRAAKLQDCTMLVCGDDLVVICSAGTQBDASIRAFTEAMTRY 1740  
 2706 TSCGNLTCTYLKAAACRAAKLQDCTMLVCGDDLVVICSAGTQBDASIRAFTEAMTRY 2765  
 1741 SAPPDGPPEYDLELITSCSSNVSAHDASGRVYVLTFRDPTTPLARAWEETARHTPVN 1800  
 2766 SAPPDGPPEYDLELITSCSSNVSAHDASGRVYVLTFRDPTTPLARAWEETARHTPVN 2825  
 1801 SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIQRL 1860  
 2826 SWLGNIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIQRL 2885  
 1861 HGLSAFSLHSYSGENRVASCLRLKGVPEPLRVWRHRSVRARLLSQGGRATCGKLYF 1920  
 2886 HGLSAFSLHSYSGENRVASCLRLKGVPEPLRVWRHRSVRARLLSQGGRATCGKLYF 2945  
 1921 NNAVRTKLTLPAAASQDLSSWVAGYSGGDIYHLSRARPWFMCWLLLSVGVGIY 1980  
 2946 NNAVRTKLTLPAAASQDLSSWVAGYSGGDIYHLSRARPWFMCWLLLSVGVGIY 3005  
 1981 LLPNR 1985  
 3006 LLPNR 3010

RESULT 3

Q9J3H2  
 ID Q9J3H2 PRELIMINARY; PRT; 3010 AA.  
 AC Q9J3H2;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyprotein.  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 CC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=MD20;  
 RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF207761; AAF65951.1;  
 DR PIR: A61196; A61196.  
 DR PIR: PQ0246; PQ0246.  
 DR HSSP: P26663; 1QUV.  
 DR GO: GO:0016021; C:integral to membrane; IEA.  
 DR GO: GO:0019028; C:viral capsid; IEA.  
 DR GO: GO:0019031; C:viral envelope; IEA.  
 DR GO: GO:0005824; F:ATP binding; IEA.  
 DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.  
 DR GO: GO:0003723; F:RNA binding; IEA.  
 DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
 DR GO: GO:0008236; F:serine-type peptidase activity; IEA.  
 DR GO: GO:0005198; F:structural molecule activity; IEA.  
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR GO: GO:0006350; P:transcription; IEA.  
 DR GO: GO:0019079; P:viral genome replication; IEA.  
 DR GO: GO:0019087; P:viral transformation; IEA.  
 DR InterPro: IPR000345; CytC heme\_BS.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR002522; HCV capsid.  
 DR InterPro: IPR002521; HCV core.  
 DR InterPro: IPR002519; HCV env.  
 DR InterPro: IPR002531; HCV NS1.  
 DR InterPro: IPR000745; HCV NS4a.  
 DR InterPro: IPR001490; HCV NS4b.  
 DR InterPro: IPR002868; HCV NS5a.  
 DR InterPro: IPR002166; HCV RdRp.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR004109; Peptidase\_S29.

InterPro; IPR009003; Pept. Ser. Cys.  
DR InterPro; IPR002518; Pept\_U39\_HCV\_Ns2.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR007094; RNA\_pol\_DS\_PSVir.  
DR Pfam; PF01543; HCV\_capsid; 1.  
DR Pfam; PF01542; HCV\_core; 1.  
DR Pfam; PF01539; HCV\_env; 1.  
DR Pfam; PF01560; HCV\_NS1; 1.  
DR Pfam; PF01538; HCV\_NS2; 1.  
DR Pfam; PF02907; HCV\_NS3; 1.  
DR Pfam; PF01006; HCV\_NS4a; 1.  
DR Pfam; PF01001; HCV\_NS4b; 1.  
DR Pfam; PF01506; HCV\_NS5a; 1.  
DR Pfam; PF00271; Helicase\_C; 1.  
DR Pfam; PF00998; Viral\_RdRP; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
SQ SEQUENCE 3010 AA; 326764 MW; 1A48EB4B5E1440D0 CRC64;  
  
Query Match 97.9%; Score 10245; DB 2; Length 3010;  
Best Local Similarity 97.2%; Pred. No. 0;  
Matches 1929; Conservative 30; Mismatches 26; Indels 0; Gaps 0;  
  
QY 1 MAPTAYSQOTRGLGCIITSLTGRDRNVEGEVQVYVSTATQSLATCVNGVCWTVYHGA 60  
DB 1026 LAPITATQOTRGLGCIITSLTGRDRNVEGEVQVYVSTATQSLATCVNGVCWTVYHGA 1085  
  
QY 61 GSKTLGAPKFIITQMYTNVQDLVGMQAPPGARSLTPTCTGSSDLYLVTRHADVIPVRRR 120  
DB 1086 GSKALAGQKPVQMYTNVQDLVGMQAPPGARSLTPTCTGSSDLYLVTRHADVIPVRRR 1145  
  
QY 121 GDSRGLSPRPVSYLKGSSGGPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
DB 1146 GDSRGLSPRPVSYLKGSSGGPLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205  
  
QY 181 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSGKSTKVPAAYAAQGVKVLVNPVAAATLFGF 240  
DB 1206 RSPVFTDNSSPPAVPOTFOVAHLHAPTGSGKSTKVPAAYAAQGVKVLVNPVAAATLFGF 1265  
  
QY 241 AYMSKAGHDIPNIRTGVRTITGAPITYSYGKFLADGCGSGGAYDIIICDECHSTDSTT 300  
DB 1266 AYMSKAGHDIPNIRTGVRTITGAPITYSYGKFLADGCGSGGAYDIIICDECHSTDSTT 1325  
  
QY 301 ILGIGTVLDOAETAGARLVVLAATPPGSVTVPHNPTEEVALSSTGPIPYGKAIPETI 360  
DB 1326 ILGIGTVLDOAETAGARLVVLAATPPGSVTVPHNPTEEVALSSTGPIPYGKAIPETI 1385  
  
QY 361 KGRGHILFCHSKKKKDELAALKSLGLNAVAYVYRGLDVSIPTSGDVIIVATDALMTGFT 420  
DB 1386 KGRGHILFCHSKKKKDELAALKSLGLNAVAYVYRGLDVSIPTSGDVIIVATDALMTGFT 1445  
  
QY 421 GDFSVIDNCTVTQVDFSLDFTFTIETTTPQDAVRSQRGRTRGRGMGIYRFVTPG 480  
DB 1446 GDFSVIDNCTVTQVDFSLDFTFTIETTTPQDAVRSQRGRTRGRGMGIYRFVTPG 1505  
  
QY 481 ERPSGMFDSVLCEDYDAGCANYELTPAETSRLRAYLNTPLPVCODHLEFWESVFTGL 540  
DB 1506 ERPSGMFDSVLCEDYDAGCANYELTPAETSRLRAYLNTPLPVCODHLEFWESVFTGL 1565  
  
QY 541 THIDAHFLSOTKAGDNFPLVAYQATVCARAQAPPSWQMKCLIRLXPTLHGTPPLL 600  
DB 1566 THIDAHFLSOTKAGDNFPLVAYQATVCARAQAPPSWQMKCLIRLXPTLHGTPPLL 1625  
  
QY 601 YRLGAVQNEVTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAYCLTITGSSVIVGR 660  
DB 1626 YRLGAVQNEVTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAYCLTITGSSVIVGR 1685  
  
QY 661 IILSGKPAITPDREVLRYEFDEMEBCASHLPYIEQGMQLAEQFKQKAIGLLQATKQAEA 720  
DB 1686 IILSGKPAVTPDRDLVLYQEFDEMEBCASHLPYIEQGMQLAEQFKQKAIGLLQATKQAEA 1745

QY 721 AAPVVESKWTLEAPFAWKHMWNFIISGIOYLAGLSTLFCNPAIASIMAFPTASITSLPTTQH 780  
DB 1746 AAPVVESKWALETFWAKHMWNFIISGIOYLAGLSTLFCNPAIASIMAFPTASITSLPTTQH 1805  
  
QY 781 TLLFNILGGWVAQAQAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840  
DB 1806 TLLFNILGGWVAQAQAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 1865  
  
QY 841 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHHVGPGEVAVQMMNRLIAFASR 900  
DB 1866 FKVMGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHHVGPGEVAVQMMNRLIAFASR 1925  
  
QY 901 GNVSPHYYPESDAAARVTQILSSLTITQLLKHOWINEDCSTPCSGSWLRDWDWIC 960  
DB 1926 GNVSPHYYPESDAAARVTQILSSLTITQLLKHOWINEDCSTPCSGSWLRDWDWIC 1985  
  
QY 961 TVLTDFTKWLQSKLLPRLPGVPFFSCORGKYGVMRGDGMOTTCPCGAOITGHVKNCSMR 1020  
DB 1986 AVLTDFTKWLQSKLLPRLPGVPFFSCORGKYGVMRGDGMOTTCPCGAOITGHVKNCSMR 2045  
  
QY 1021 IVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGFHYVTGM 1080  
DB 2046 IVGPRTCSNTWHGTFPINAYTTGCTPSPAPNYSRALMRVAABEYVEVTRVGFHYVTGM 2105  
  
QY 1081 TTDNVKCPQVPAPEPTEVDGVLHRYAPACKPLLBEEVTFVLGLNQYLVSQLPCEPE 1140  
DB 2106 TTDNVKCPQVPAPEPTEVDGVLHRYAPACKPLLBEEVTFVLGLNQYLVSQLPCEPE 2165  
  
QY 1141 PDVAVLTSMLTDSHITAEAKRLARGSPSPSLASSASQSLAPSLKATCTTRHSDPAD 1200  
DB 2166 PDVAVLTSMLTDSHITAEAKRLARGSPSPSLASSASQSLAPSLKATCTTRHSDPAD 2225  
  
QY 1201 LIPANLLWRQEMGNI TRVESENKWLDSFEPLQAEDEEREVSVAEILRKRKFFRAM 1260  
DB 2226 LIPANLLWRQEMGNI TRVESENKWLDSFEPLQAEDEEREVSVAEILRKRKFFRAM 2285  
  
QY 1261 PIWARPDPNPLESKKDPDYVPVHGCPLPAKAPPPIPPRRKRITVLSESTVSALA 1320  
DB 2286 PIWARPDPNPLESKKDPDYVPVHGCPLPAKAPPPIPPRRKRITVLSESTVSALA 2345  
  
QY 1321 ELATKTFGSESAVDSGTATSDPQSDPDGSDGSDVESYSSMPLEGEPEGDPDLSGWS 1380  
DB 2346 ELATKTFGSESAVDSGTATSDPQSDGSDGSDVESYSSMPLEGEPEGDPDLSGWS 2405  
  
QY 1381 STVSEBASEDVVCCSMSTWTGALITPCAAEETKLIPINALSNLLRHHNIVYATTSRSAS 1440  
DB 2406 STVSEBASEDVVCCSMSTWTGALITPCAAEETKLIPINALSNLLRHHNIVYATTSRSAS 2465  
  
QY 1441 LROKKYTFDELQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
DB 2466 LROKKYTFDELQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 2525  
  
QY 1501 DVNLSKSAVNHIRSVMKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDLG 1560  
DB 2526 DVNLSKSAVNHIRSVMKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVPDLG 2585  
  
QY 1561 VRVCEKVALYDVVSTLPQAVMGSSYGFQYSPQGVVEFLVNAWKAKCPMGFAYDTRCFDS 1620  
DB 2586 VRVCEKVALYDVVSTLPQAVMGSSYGFQYSPQGVVEFLVNAWKAKCPMGFAYDTRCFDS 2645  
  
QY 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGNCGYRRRCRASGVL 1680  
DB 2646 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYTGGPLTNSKGNCGYRRRCRASGVL 2705  
  
QY 1681 TSCGNTLTCLYKAAAACRAAKLQDCDMLVCGDDLVIICESAGTOEDEASIRAFETAMTRY 1740  
DB 2706 TSCGNTLTCLYKAAAACRAAKLQDCDMLVCGDDLVIICESAGTOEDEASIRAFETAMTRY 2765  
  
QY 1741 SAPPDGPPEYDLELITSCSSNVSVAHADAGSKRVYLLTRDPTTPIARAAMETARITPVN 1800  
DB 2766 SAPPDGPPEYDLELITSCSSNVSVAHADAGSKRVYLLTRDPTTPIARAAMETARITPVN 2825  
  
QY 1801 SWLGNLIINTAPTILWARMILMTHFFSILLAQEQLEKALDQCIYGACYSIEPLDLPQIIQRL 1860

```
Db 2826 SWLGNIIYAPLWARMIMTHFFSILLAQEKLDCQIYGACYSIEPLDLPQIQL 2885
QY 1861 HGLSAFSLHSYSPGBINRVASCLRLKLGVPPLRWRHRARSVRARLLSQGGRATCGKLYF 1920
Db 2886 HGLSAFSLHSYSPGBINRVASCLRLKLGVPPLRWRHRARSVRARLLSQGGRATCGKLYF 2945
QY 1921 NNAVTKLTLTIPPAASQLDLSWFWAGYSGGDIYHLSRARPFRWMCILLLSVGVGY 1980
Db 2946 NNAVTKLTLTIPPAASHLDLSWFWAGYSGGDIYHLSRARPFRWMCILLLSVGVGY 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010

RESULT 4
Q9J3F4
ID Q9J3F4 PRELIMINARY; PRT; 3008 AA.
AC Q9J3F4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
GN Name=MD34;
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD34;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF208024; AAF61205.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PS0329; PS0329.
DR HSP; P26663; IJXP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006350; P:transcription; IEA.
DR GO; GO:0019079; P:viral genome replication; IEA.
DR GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; Cyt c heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRP.
DR InterPro; IPR001650; Helicase C.
DR InterPro; IPR004109; Peptidase S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_PSvir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
```

```
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase C; 1.
DR Pfam; PF00998; Viral RdRP; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN_1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3008 AA; 326835 MW; 99AE09E14C3109F4 CRC64;

Query Match 97.7%; Score 10225; DB 2; Length 3008;
Best Local Similarity 97.1%; Pred No. 0;
Matches 1928; Conservative 33; Mismatches 24; Indels 0; Gaps 0;

QY 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVSTATQSFILATCNGVCWTYYHGA 60
Db 1024 LAPITAYSQOTRGLGCIITSLTGRDRKNQVEGEVQVSTATQSFILATCNGVCWTYYHGA 1083
QY 61 GSKTLAGPKGPTOMYTNVDQDLVGWOAPPGARSITPCTCGSSDLYLVTRHADVIPVRR 120
Db 1084 GSKTLAGPKGPTOMYTNVDQDLVGWOAPPGARSITPCTCGSSDLYLVTRHADVIPVRR 1143
QY 121 GDSRGLLSRPVSVYLGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180
Db 1144 GDSRGLLSRPVSVYLGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1203
QY 181 RSPVFTDNSSPPAVPQTQVAHLHAPTGGSKTKVPAAYAAQGYKVVLNPSVAATLPGF 240
Db 1204 RSPVFTDNSSPPAVPQTQVAHLHAPTGGSKTKVPAAYAAQGYKVVLNPSVAATLPGF 1263
QY 241 AYMSKAHGIDNRITGVRTITTTGAPTISTYCKELADGCGSGAVDIIICDCHSDSTT 300
Db 1264 AYMSKAHGIDNRITGVRTITTTGAPTISTYCKELADGCGSGAVDIIICDCHSDSTT 1323
QY 301 ILGIGTVLDAQETAGARLVILATATAPPGSVTVPHENIEVALSSTGEIPFYKAIPIETI 360
Db 1324 ILGIGTVLDAQETAGARLVILATATAPPGSVTVPHENIEVALSSTGEIPFYKAIPIETI 1383
QY 361 KGRHLIFCHSKKKCDLAALKSLGLNNAVYRGLDVSVIPTSGDIVVATDALTMTGTT 420
Db 1384 KGRHLIFCHSKKKCDLAALKSLGLNNAVYRGLDVSVIPTSGDIVVATDALTMTGTT 1443
QY 421 GDFDSVIDCNTCVTQVDFSLDPTTETTTTTPQDAVSRSORRGRTGRGMGIYRVTTPG 480
Db 1444 GDFDSVIDCNTCVTQVDFSLDPTTETTTTTPQDAVSRSORRGRTGRGMGIYRVTTPG 1503
QY 481 ERPSGMFSSVLCECYDAGCANVELTPAETS VRLRAYLNTPLPGVQDHLFEWESVFTGL 540
Db 1504 ERPSGMFSSVLCECYDAGCANVELTPAETS VRLRAYLNTPLPGVQDHLFEWESVFTGL 1563
QY 541 THIDAHFISQTKOAGDNFPYLVAQATVCARQAQPPSDWDQWKCLIRLKPITHGTPPLL 600
Db 1564 THIDAHFISQTKOAGDNFPYLVAQATVCARQAQPPSDWDQWKCLIRLKPITHGTPPLL 1623
QY 601 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVTSWLVGVGVAALAAAYCLTTSVVIVGR 660
Db 1624 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVTSWLVGVGVAALAAAYCLTTSVVIVGR 1683
QY 661 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKKAIQLLOTATKQABA 720
Db 1684 IILSGKPAIIPDREVLVREFDEMEECASHLPYIEQGMQLAEQFKKAIQLLOTATKQAEV 1743
QY 721 AAPVVEKWRITLEAFWAKHMWNFISGIQYLAGISTILPGNPAIASIMAFASITSPITTOH 780
Db 1744 AAPVVEKWRITLEAFWAKHMWNFISGIQYLAGISTILPGNPAIASIMAFASITSPITTOH 1803
QY 781 TLLFNILGWAAQLAPPSSAASAFVAGIAGAVSGISGLGKVLVDILAGYAGAGALVA 840
Db 1804 TLLFNILGWAAQLAPPSSAASAFVAGIAGAVSGISGLGKVLVDILAGYAGAGALVA 1863
QY 841 FKVMGEMPSTEDLVNLLPAILSPGALVVGVVCAAILRRHVGEVAVOMNRLIAPASR 900
```

1864 FKVMSGEMSTEDLVNLLPAILSPGLVGVVCAILLRRVVGFGGAVQWNNRLLIAPSR 1923  
901 GNVHSTHYVPESDAAARVTOILSSLTITQLLKLHQMINEDCSTPCSGSWLRDVMWDMIC 960  
1924 GNVHSTHYVPESDAAARVTOILSSLTITQLLKLHQMINEDCSTPCSGSWLRDVMWDMIC 1983  
961 TVLTDFKTLWLOSLLRLPGVPFSCORGKYGKVRGDMQMTTCPCGAQITGHVKGSMR 1020  
1984 TVLTDFKTLWLOSLLRLPGVPFSCORGKYGKVRGDMQMTTCPCGAQITGHVKGSMR 2043  
1021 IUGPRCTSNTHGCTFPINAYTTCPTSPAPNYSRALRWVAABEYEVTVRGDFHYVTGM 1080  
2044 IUGPKCTSNTHGCTFPINAYTTCPTSPAPNYSRALRWVAABEYEVTVRGDFHYVTGM 2103  
1081 TTDNVKCPQVPAPEFTEVDGVRLLHRYAPACKPLLRREEVTVLGLNQYLVGSQLPCEPE 1140  
2104 TTDNVKCPQVPAPEFTEVDGVRLLHRYAPACKPLLRREEVTVLGLNQYLVGSQLPCEPE 2163  
1141 POVAVLTSMLTDSHITAEAKRLARGSPSPSLASSASQLSAPSLKATCTTTHDSDPAD 1200  
2164 POVAVLTSMLTDSHITAEAKRLARGSPSPSLASSASQLSAPSLKATCTTTHDSDPAD 2223  
1201 LIEANLLWRQEMGNGNITRVESENKVVILDSFELQAEDEDEREVSVPAILLRKRKPPRAM 1260  
2224 LIEANLLWRQEMGNGNITRVESENKVVILDSFELQAEDEDEREVSVPAILLRKRKPPRAM 2283  
1261 PIWARDYNPPLLESWKDPDYPVHVHGCPLPPAKAPPIPPRRKRKTVLSESTVSALA 1320  
2284 PIWARDYNPPLLESWKDPDYPVHVHGCPLPPAKAPPIPPRRKRKTVLSESTVSALA 2343  
1321 ELATKTFGSESAVDSGTATASPDQSDSDGSDVESYSSMPPELGEFGDPLSDGSW 1380  
2344 ELATKTFGSESAVDSGTATAPDQSDSDGSDVESYSSMPPELGEFGDPLSDGSW 2403  
1381 STVSEASEDVCVCCMSYTWGTALITPCAAETKLPINALNSLLRHHNLVYATTSRSAS 1440  
2404 STVSEASEDVCVCCMSYTWGTALITPCAAEGSKLPINALNSLLRHHNLVYATTSRSAS 2463  
1441 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFGYGA 1500  
2464 LRQKVTDFRLQVLDHRYDVLKEMKAKASTVKALLSVEEACKLTPPHSARSKFGYGA 2523  
1501 DVNLSKAVNHRISVWKOLLETEPIDITIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560  
2524 DVNLSKAVNHRISVWKOLLETEPIDITIMAKNEIFCVQPEKGRKPARLIVFPDLG 2583  
1561 VRVCERWALYDVVSTLPQAVMGSSYGFQYSPGQVVEFLVNAWKAKCPMGFAYDTRCFDS 1620  
2584 VRVCERWALYDVVSTLPQAVMGSSYGFQYSPGQVVEFLVNAWKAKCPMGFAYDTRCFDS 2643  
1621 TVTENDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPVTNSKQNGCYRRCRASGVLT 1680  
2644 TVTESDIRVEESTYQCCDLAPEARQAIRSLTERLYTGGPVTNSKQNGCYRRCRASGVLT 2703  
1681 TSCGNLTLYLKAAACRAKLODCTMLVCGDDLVI CESAGTQDEASIRATEAMTRY 1740  
2704 TSCGNLTLYLKAAACRAKLODCTMLVCGDDLVI CESAGTQDEASIRATEAMTRY 2763  
1741 SAPPDGPPEYDELELITSCSNVSVAHADSGKRVYVLTDRPTTPLARAWEATARTPVN 1800  
2764 SAPPDGPPEYDELELITSCSNVSVAHADSGKRVYVLTDRPTTPLARAWEATARTPVN 2823  
1801 SWLGNITMVAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPQIIQRL 1860  
2824 SWLGNITMVAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPQIIQRL 2883  
1861 HGLSASLSHSYSCEINRVASCLRKLGVPPLRVHRARSVRARLLSQGGRATCGKLYF 1920  
2884 HGLSASLSHSYSCEINRVASCLRKLGVPPLRVHRARSVRARLLSQGGRATCGKLYF 2943  
1921 NNAVRTKLTPTIPAAASQDLSNFWAGYGGDIYHLSLRARPRWFMCLLLLSVGVGIY 1980

Db 2944 NNAVRTKLTPTIPAAASQDLSNFWAGYGGDIYHLSLRARPRWFMCLLLLSVGVGIY 3003  
QY 1981 LLENR 1985  
Db 3004 LLENR 3008  
RESULT 5  
Q8V638 PRELIMINARY; PRT; 3010 AA.  
AC Q8V638;  
DT 01-MAR-2002 (T-EMBLrel. 20, Created)  
DT 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)  
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)  
DE Polyprotein (Fragment).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Fanning L.J., Itakura J., Nagayama K., Enomoto N.;  
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF313916; AAL55821.1;  
DR PIR: A61196; A61196.  
DR PIR: PQ0246; PQ0246.  
DR PIR: PQ0804; PQ0804.  
DR PIR: PS0329; PS0329.  
DR GO: GO:0016021; C:integral to membrane; IEA.  
DR GO: GO:0019028; C:viral capsid; IEA.  
DR GO: GO:0019031; C:viral envelope; IEA.  
DR GO: GO:0005524; F:ATP binding; IEA.  
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO: GO:0003723; F:RNA binding; IEA.  
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO: GO:0005198; F:structural molecule activity; IEA.  
DR GO: GO:0006508; F:proteolysis and peptidolysis; IEA.  
DR GO: GO:0006350; P:transcription; IEA.  
DR GO: GO:0019079; P:viral genome replication; IEA.  
DR GO: GO:0019087; P:viral transformation; IEA.  
DR InterPro: IPR000345; CytC\_heme\_BS.  
DR InterPro: IPR001410; DEAD.  
DR InterPro: IPR002522; HCV\_capsid.  
DR InterPro: IPR002521; HCV\_core.  
DR InterPro: IPR002519; HCV\_env.  
DR InterPro: IPR002531; HCV\_NS1.  
DR InterPro: IPR000745; HCV\_NS4a.  
DR InterPro: IPR001490; HCV\_NS4b.  
DR InterPro: IPR002868; HCV\_NS5a.  
DR InterPro: IPR002166; HCV\_RdRP.  
DR InterPro: IPR001650; Helicase\_C.  
DR InterPro: IPR004109; Peptidase\_S29.  
DR InterPro: IPR009003; Pept\_Ser\_Cys.  
DR InterPro: IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR Pfam: PF01543; HCV\_capsid; 1.  
DR Pfam: PF01542; HCV\_core; 1.  
DR Pfam: PF01539; HCV\_env; 1.  
DR Pfam: PF01538; HCV\_NS1; 1.  
DR Pfam: PF01560; HCV\_NS2; 1.  
DR Pfam: PF02907; HCV\_NS3; 1.  
DR Pfam: PF01006; HCV\_NS4a; 1.  
DR Pfam: PF01001; HCV\_NS4b; 1.  
DR Pfam: PF01506; HCV\_NS5a; 1.  
DR Pfam: PF00271; Helicase\_C; 1.  
DR SMART: SM00487; DEXD; 1.  
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
FT NON\_TER 3010 3010

SQ	SEQUENCE	3010 AA; 327182 MW; 33AAA6C07251C839 CRC64;	
	Query Match	97.7%; Score 10225; DB 2; Length 3010;	
	Best Local Similarity	97.2%; Pred. No. 0;	
	Matches 1930; Conservative	29; Mismatches 26; Indels 0; Gaps 0;	
QY	1	MAPITAYSQOTRGLGCIITSITGRDRNOVEGEVWVSTATOSFLATCVNGVCWTVYHGA 60	
DB	1026	LAPITAYAQTRGLGCIITSITGRDRNOVEGEVWVSTATOSFLATCVNGVCWTVYHGA 1085	
QY	61	GSKTLGPKGPIQOMYNTVNDQVLGVQWQAPPGARSUTPTCCGSSDLYLVTRHADVIVRRR 120	
DB	1086	GKTLGAGKGPITOMYNTVNDQVLGVWLAAPPARSUTPTCCGSSDLYLVTRHADVIVRRR 1145	
QY	121	GDSRGLSPRVSVYLKSSGGPILCPSCGHAVGIFRAAVCTRGVAKAVDFVSVESMETTM 180	
DB	1146	GDSRGLSPRVSVYLKSSGGPILCPSCGHAVGIFRAAVCTRGVAKAVDFVSVESMETTM 1205	
QY	181	RSPVFTDNSSPPAVPQTFOVAHLHAPTGSCKTKVPAAYAAQGYKVLVNLNPSVAATLGF 240	
DB	1206	RSPVFTDNSSPPAVPQTFOVAHLHAPTGSCKTKVPAAYAAQGYKVLVNLNPSVAATLGF 1265	
QY	241	AYMSKAHGDIPNIRTVGVRITITGAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 300	
DB	1266	AYMSKAHGDIPNIRTVGVRITITGAPITYSTYKFLADGGCGGAYDIIICDECHSTDST 1325	
QY	301	ILGIGTVLDOAETAGARLVLATATPPGSVTVPHNIEVALSSTGEIPFYKKAIPETI 360	
DB	1326	ILGIGTVLDOAETAGARLVLATATPPGSVTVPHNIEVALSSTGEIPFYKKAIPETI 1385	
QY	361	KGGRHLIFCHSKKKCKDELAALKSGLGNVAVYRGLDVSVIPTSGDVIIVATDALMTGT 420	
DB	1386	KGGRHLIFCHSKKKCKDELAALKSGLGNVAVYRGLDVSVIPTSGDVIIVATDALMTGT 1445	
QY	421	GDPSVIDCNTVOTQVDFSLDPTFTIETTVPQDAVSQRSQRRGTGRGMGIYFVPTG 480	
DB	1446	GDPSVIDCNTVOTQVDFSLDPTFTIETTVPQDAVSQRSQRRGTGRGMGIYFVPTG 1505	
QY	481	EPSGMFDSSVLCECYDAGCAMVELTPAETSURLRAYLNTGILPVQODHLEWESVFTGL 540	
DB	1506	EPSGMFDSSVLCECYDAGCAMVELTPAETSURLRAYLNTGILPVQODHLEWESVFTGL 1565	
QY	541	THIDAHFLSQTQAGDNFYLVAQATVCARAQAPPPSDQMWKCLIRLKPILHGTPLL 600	
DB	1566	THIDAHFLSQTQAGDNFYLVAQATVCARAQAPPPSDQMWKCLIRLKPILHGTPLL 1625	
QY	601	YRLGAVQNEVTHPTIKTIFIMACMSADLEWVTSWVLVGGVLAALAAAYCLITGSSVIVGR 660	
DB	1626	YRLGAVQNEVTHPTIKTIFIMACMSADLEWVTSWVLVGGVLAALAAAYCLITGSSVIVGR 1685	
QY	661	IILSGKPAIIPREVLYRSEFDEMEECASHLPVTEQMOQLAQOPKOKAIGLLOTATKQAEA 720	
DB	1686	IILSGKPAIIPREVLYRSEFDEMEECASHLPVTEQMOQLAQOPKOKAIGLLOTATKQAEA 1745	
QY	721	AAPVVEKWRITLEAFWAKEMWNPISGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOH 780	
DB	1746	AAPVVEKWRITLEAFWAKEMWNPISGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOH 1805	
QY	781	TLLFNILGWAAQIAPPAAAFVAGIAGAAVSGIGLKVLDIILAGYAGAGALVA 840	
DB	1806	TLLFNILGWAAQIAPPAAAFVAGIAGAAVSGIGLKVLDIILAGYAGAGALVA 1865	
QY	841	EKVMGEMSTEDLVNLLPAILSPGALVGVVCAATILRRHVGPGEAGVQMMRLIAFASR 900	
DB	1866	EKVMGEMSTEDLVNLLPAILSPGALVGVVCAATILRRHVGPGEAGVQMMRLIAFASR 1925	
QY	901	GNHVSPTHVPESDAAARVTQILSSLTITQLKRLHWINEDCSTPCSGSWLRDWDWMC 960	
DB	1926	GNHVSPTHVPESDAAARVTQILSSLTITQLKRLHWINEDCSTPCSGSWLRDWDWMC 1985	
QY	961	TVLTDFTKWLQSKLLPRLFGVPPFCQGYKGYWRGDMQITCPCGAQITGHVKNQSMR 1020	
DB	1986	TVLTDFTKWLQSKLLPRLFGVPPFCQGYKGYWRGDMQITCPCGAQITGHVKNQSMR 2045	

QY	1021	IVGPETCSNTWHGTTPINAYTTGCTPSPAPNYSRALMRVAEEYVEVTRVGDHFHYTCM 1080	
DB	2046	IVGPETCSNTWHGTTPINAYTTGCTPSPAPNYSRALMRVRAEEYVEVTRVGDHFHYTCM 2105	
QY	1081	TTDNYKCPQCPAPPEFFTEVDGVRHLRYAPACKPLREBEVTVLGNLQYLVGSQPCBPE 1140	
DB	2106	TTDNYKCPQCPAPPEFFTEVDGVRHLRYAPACKPLREBEVTVLGNLQYLVGSQPCBPE 2165	
QY	1141	PDVAVLTSMLTDPSSHITAEATAKRLARGSPSPSLASSASQLSAPSLKATCTTRHDSPAD 1200	
DB	2166	PDVAVLTSMLTDPSSHITAEATAKRLARGSPSPSLASSASQLSAPSLKATCTTRHDSPAD 2225	
QY	1201	LIEANLLWRQEMGNIITRVESENKVVILDSPEPLQAEEDEREVSYPAEILRRSRKPPRAM 1260	
DB	2226	LIEANLLWRQEMGNIITRVESENKVVILDSPEPLQAEEDEREVSYPAEILRRSRKPPRAM 2285	
QY	1261	PIWARPDYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSSTVSSALA 1320	
DB	2286	PIWARPDYNPPLLESWKDPDYVPPVHGCPLPPAKAPPIPPRRKRTVVLSSTVSSALA 2345	
QY	1321	ELATKTFGSSSESAVDSGTATASPOPSDDGDAGSDVESYSSMPLECEPGDPLSDGSM 1380	
DB	2346	ELATKTFGSSSESAVDSGTATASPOPSDDGDAGSDVESYSSMPLECEPGDPLSDGSM 2405	
QY	1381	STVSEEAAGEDVVCSSMSYTWGALITPCAABETKLPINALSNLLRHHNLVYATTSSRSAS 1440	
DB	2406	STVSEEAAGEDVVCSSMSYTWGALITPCAABETKLPINALSNLLRHHNLVYATTSSRSAS 2465	
QY	1441	LQKQVTVFDRLOVLDHYRDVLKEMKAKASTYKAKLLSVEEACKLTPPHSARSKFGYGA 1500	
DB	2466	LQKQVTVFDRLOVLDHYRDVLKEMKAKASTYKAKLLSVEEACKLTPPHSARSKFGYGA 2525	
QY	1501	DYRNLSSKAVAHIRSVWKDLLEDTPIDTTIMAKNEVFCVQPEKGRKPAELIVFPDLG 1560	
DB	2526	DYRNLSSKAVAHIRSVWKDLLEDTPIDTTIMAKNEVFCVQPEKGRKPAELIVFPDLG 2585	
QY	1561	VVCEKMAIYDVVSTLPOAVMGSSYGVQYSPQORVEFLVNAWKAKKCPMGFYDTRCFDS 1620	
DB	2586	VVCEKMAIYDVVSTLPOAVMGSSYGVQYSPQORVEFLVNAWKAKKCPMGFYDTRCFDS 2645	
QY	1621	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRCRAGSVLT 1680	
DB	2646	TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRCRAGSVLT 2705	
QY	1681	TSCGNTLTCYLKAAACRAAKLODCTMLVCGDLVVICESAGTQDEASLRAFTAMTRY 1740	
DB	2706	TSCGNTLTCYLKAAACRAAKLODCTMLVCGDLVVICESAGTQDEASLRAFTAMTRY 2765	
QY	1741	SAPPGDPPKPEYDLELITSCSSNVSAHDASGRVYVYLTRDPTTTLARAAMETARHTPVN 1800	
DB	2766	SAPPGDPPKPEYDLELITSCSSNVSAHDASGRVYVYLTRDPTTTLARAAMETARHTPVN 2825	
QY	1801	SWLGNIIYVAPTLWARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIQL 1860	
DB	2826	SWLGNIIYVAPTLWARMILMTHFFSILLAQEOLEKALDCQIYGACYSIEPLDLPQIQL 2885	
QY	1861	HGLSFAFSLHSYSGEINRVASCLKLGYPPLRVWHRARSVEARLLSQGRAATCGKYL 1920	
DB	2886	HGLSFAFSLHSYSGEINRVASCLKLGYPPLRVWHRARSVEARLLSQGRAATCGKYL 2945	
QY	1921	NNAVTKLKITPIPAASQDLSSWFVAGYSGGDIYHLSRARPRWFMWMLLLLSVGVGY 1980	
DB	2946	NNAVTKLKITPIPAASQDLSSWFVAGYSGGDIYHLSRARPRWFMWMLLLLSVGVGY 3005	
QY	1981	LLPNR 1985	
DB	3006	LLPNR 3010	

RESULT 6  
O93016  
ID O93016 PRELIMINARY; PRT; 3010 AA.



AC O93016;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Polyprotein.  
 OS Hepatitis C virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
 OC Hepacivirus.  
 OX NCBI\_TaxID=11103;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98232263; PubMed=9572551;  
 RA Trowbridge R., Gowans E.J.;  
 RT "Molecular cloning of an Australian isolate of hepatitis C virus.";  
 RL Arch. Virol. 143:501-511(1998).  
 DR EMBL: AJ000009; CAA03854.1; -.  
 DR PIR: A61196; A61196.  
 DR PIR: P02446; P02446.  
 DR PIR: P02555; P02555.  
 DR PIR: P0329; P0329.  
 DR HSP: P26663; JUXP.  
 DR GO: GO:0016021; C:integral to membrane; IEA.  
 DR GO: GO:0019028; C:viral capsid; IEA.  
 DR GO: GO:0019031; C:viral envelope; IEA.  
 DR GO: GO:0005524; F:ATP binding; IEA.  
 DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.  
 DR GO: GO:0003723; F:RNA binding; IEA.  
 DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
 DR GO: GO:0008236; F:serine-type peptidase activity; IEA.  
 DR GO: GO:0005198; F:structural molecule activity; IEA.  
 DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.  
 DR GO: GO:0006350; P:transcription; IEA.  
 DR GO: GO:0019079; P:viral genome replication; IEA.  
 DR GO: GO:0019087; P:viral transformation; IEA.  
 DR InterPro: IPR000345; CytC\_heme\_BS.  
 DR InterPro: IPR001410; DEAD.  
 DR InterPro: IPR002522; HCV capsid.  
 DR InterPro: IPR002521; HCV\_core.  
 DR InterPro: IPR002519; HCV\_env.  
 DR InterPro: IPR002531; HCV\_NS1.  
 DR InterPro: IPR000745; HCV\_NS4a.  
 DR InterPro: IPR001490; HCV\_NS4b.  
 DR InterPro: IPR002868; HCV\_NS5a.  
 DR InterPro: IPR002166; HCV\_RdRP.  
 DR InterPro: IPR001650; Helicase\_C.  
 DR InterPro: IPR004109; Peptidase\_S29.  
 DR InterPro: IPR009003; Pept\_Ser\_Cys.  
 DR InterPro: IPR002518; Pept\_U39\_HCV\_NS2.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR007094; RNA\_pol\_PSVir.  
 DR Pfam: PF01543; HCV\_capsid; 1.  
 DR Pfam: PF01542; HCV\_core; 1.  
 DR Pfam: PF01539; HCV\_env; 1.  
 DR Pfam: PF01560; HCV\_NS1; 1.  
 DR Pfam: PF01538; HCV\_NS2; 1.  
 DR Pfam: PF02907; HCV\_NS3; 1.  
 DR Pfam: PF01006; HCV\_NS4a; 1.  
 DR Pfam: PF01001; HCV\_NS4b; 1.  
 DR Pfam: PF01506; HCV\_NS5a; 1.  
 DR Pfam: PF00271; Helicase\_C; 1.  
 DR Pfam: PF00998; Viral\_RdRP; 1.  
 DR SMART: SM00487; DEXDC; 1.  
 DR PROSITE: PS00180; CYTOCHROME C; UNKNOWN\_1.  
 KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
 Polyprotein; Transmembrane.  
 FT CHAIN 810 1026 NS2 protein.  
 FT CHAIN 1027 1657 NS3 protein.  
 FT CHAIN 1658 1711 NS4a protein.  
 FT CHAIN 1 191 core protein.  
 FT CHAIN 1712 1972 NS4b protein.  
 FT CHAIN 1973 2419 NS5a protein.  
 FT CHAIN 2420 3010 NS5b protein.  
 FT CHAIN 192 383 E1 protein.

FT	CHAIN	384	746	E2 protein.
FT	CHAIN	747	809	p7 protein.
SQ	SEQUENCE	3010 AA;	327121 MW;	OEE02EDA54A8B61D CRC64;

  

Query Match	97.6%;	Score 10218;	DB 2;	Length 3010;
Best Local Similarity	97.0%;	Pred. No. 0;		
Matches 1925;	Conservative 35;	Mismatches 25;	Indels 0;	Gaps 0;

  

QY	1	MAPITAYSQQTGLGCGIITSLTGRDRNQVEGEVQVWVSTATQSFATFCVNGVCWTVVHGA	60
Db	1026	LAPITAYSQQTGLGCGIITSLTGRDRNQVEGEVQVWVSTATQSFATFCVNGVCWTVVHGA	1085
QY	61	GSKTLAGPKGPIOMYTNVDQDLVGHQAPPGARSLPCTCGSSDLXVTRHADVIPVRRR	120
Db	1086	GSKTLAGPKGPVTOMYTNVDQDLVGHQAPPGARSLPCTCGSSDLXVTRHADVIPVRRR	1145
QY	121	GDGRGSLSPRPVSYLKGSGGPGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM	180
Db	1146	GDGRGSLSPRPVSYLKGSGGPGLLCPGSHAVGIFRAAVCTRGVAKAVDPVPVSMETTM	1205
QY	181	RSPVFTDNGSPPAVPQTFQVAHLHAPTGSCKSTKVPAAVAAQGYKVLVLPNSVAATLFGF	240
Db	1206	RSPVFTDNGSPPAVPQTFQVAHLHAPTGSCKSTKVPAAVAAQGYKVLVLPNSVAATLFGF	1265
QY	241	AYMSKAHGIDPNIRTVGVRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT	300
Db	1266	AYMSKAHGIDPNIRTVGVRITTCAPITYSTYKFLADGGCGGAYDIIICDECHSTDSTT	1325
QY	301	ILGIGTVLQDAETAGARLVVLTATPPGVSVPVPHNIEEVALSSTGEIPIYGAIPETI	360
Db	1326	ILGIGTVLQDAETAGARLVVLTATPPGVSVPVPHNIEEVALSSTGEIPIYGAIPETI	1385
QY	361	KGGRHLIFCHSKKKKDELAALKSLGLNNAVYVRGLDVSVPITSGDVIWVATDALMTGFT	420
Db	1386	KGGRHLIFCHSKKKKDELAALKSLGLNNAVYVRGLDVSVPITSGDVIWVATDALMTGFT	1445
QY	421	GDPSVIDCNTCVTVVDSLDPTFTIETTVQDAVSRSORGRTRGRMGYRFPVTPG	480
Db	1446	GDPSVIDCNTCVTVVDSLDPTFTIETTVQDAVSRSORGRTRGRMGYRFPVTPG	1505
QY	481	ERPSGMFDSVLCEDYDAGCAWYELTPAETSVRRLAYLNTPLGVCDDHLEWESVFTGL	540
Db	1506	ERPSGMFDSVLCEDYDAGCAWYELTPAETSVRRLAYLNTPLGVCDDHLEWESVFTGL	1565
QY	541	THIDAFLSQTKOAGDNFFYLVAQATVCARAQAPPPSWDMQWKILIRLKPTLHGPTLL	600
Db	1566	THIDAFLSQTKOAGDNFFYLVAQATVCARAQAPPPSWDMQWKILIRLKPTLHGPTLL	1625
QY	601	YRLGAVONEVTTTHPIKYMCMASADLEWVSTWLVGGVLAALAAAYCLTTGSSVIVGR	660
Db	1626	YRLGAVONEVTTTHPIKYMCMASADLEWVSTWLVGGVLAALAAAYCLTTGSSVIVGR	1685
QY	661	IILSGKPAIIPREVLYREFDEMEECASHLPYEQGMQLAEQFKOKAIGLLQATKQAEA	720
Db	1686	IILSGKPAIIPREVLYREFDEMEECASHLPYEQGMQLAEQFKOKAIGLLQATKQAEA	1745
QY	721	AAPVVESKWRTELEAFWAKHMWNFISGIQYLAGLSTLPGNPAIASLMFTAITSPLTQH	780
Db	1746	AVPVVESKWALEAFWAKHMWNFISGIQYLAGLSTLPGNPAIASLMFTAITSPLTQH	1805
QY	781	TLLFNILGGVAAQLAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA	840
Db	1806	TLLFNILGGVAAQLAPPSAASAFVAGIAGAAVSGISGLKVLVDILAGYAGVAGALVA	1865
QY	841	FKVMSEMPSTEDLVNLLPAIILSPALVGVVCAATLRRHVGPGEVAGVOMNRLTAFASR	900
Db	1866	FKVMSEMPSTEDLVNLLPAIILSPALVGVVCAATLRRHVGPGEVAGVOMNRLTAFASR	1925
QY	901	GNHVSPTHVVPESDAAARVTQILSSITITQLKRLHQWINECDSTPCSGSLRDVWDWIC	960
Db	1926	GNHVSPTHVVPESDAAARVTQILSSITITQLKRLHQWINECDSTPCSGSLRDVWDWIC	1985
QY	961	TVLITFKTLWLSKLLPLRPGVPPFFSCQGRGKGVWRGDIQMOTTCPCGCIQTHGVKNGSMR	1020

Db 1986 TVLDTFTWLSKLLPRLPGFFYSQGRYKGVWGDGIMQTTCCGQAIGCHVANGSMR 2045  
QY 1021 IVGPRTCSNTWHGTGTPINAYTGTCTPSPAPNYSRALRWVAARVVEVTRVGDFFHYVTGM 1080  
Db 2046 IVGPRTCSNTWHGTGTPINAYTGTCTPSPAPNYSRALRWVAARVVEVTRVGDFFHYVTGM 2105  
QY 1081 TTDNVKCPQVPAPPEFFTEVDGVRHRYAPACKPLREVEFLVGLNQYLVCSQPCPE 1140  
Db 2106 TTDNVKCPQVPAPPEFFTEVDGVRHRYAPACKPLREVEFLVGLNQYLVCSQPCPE 2165  
QY 1141 PDVAVLTSMLTDPDHITTAETAKRRLARGSPPSLASSASQSLAPSLKATCTTRHDSPOAD 1200  
Db 2166 PDVVVTSMLTDPDHITTAETAKRDLRGSPPSLASSASQSLAPSLKATCTTRHDSPOAD 2225  
QY 1201 LIEANLLWRQMGNNITRVESKNNVILDSFPLQAEEDERVSVPAILRRSRKFPFRAM 1260  
Db 2226 LIEANLLWRQMGNNITRVESKNNVILDSFPLQAEEDERVSVAAILRRSRKFPFRAM 2285  
QY 1261 PIWARPDPNPLLESWKDPDYVPPVHGGCLPAPKAPPIPPPRKRTVVLSSTVSSALA 1320  
Db 2286 PIWARPDPNPLLESWKDPDYVPPVHGGCLPAPKAPPIPPPRKRTVVLSSTVSSALA 2345  
QY 1321 ELATKTFGSSSAVDSGTATASDPQSDGDAGSDVESYSSMPLEGPDPDLSDGSW 1380  
Db 2346 ELATKTFGSSSAVDSGTATAPPDPDNDTGSDESCESSMPLEGPDPDLSDGSW 2405  
QY 1381 STVSEASDDVCCSMYSYTWGALITPCAABETKLPINALSNSLLRHNNLVATTSSRAS 1440  
Db 2406 STVSEASDDVCCSMYSYTWGALITPCAABESKLPINALSNSLLRHNNLVATTSSRAS 2465  
QY 1441 LQKQKVTFRLOVLDHDDHVDLKEKAKASTVKALLSVEEACKLTPPHSARSKFGYAK 1500  
Db 2466 LQKQKVTFRLOVLDHDDHVDLKEKAKASTVKALLSVEEACKLTPPHSARSKFGYAK 2525  
QY 1501 DVNRLSSKAVNHRSVWKKLLEDTEPIDTTIMAKNEFCVQPEKGRKPARLIIVFPDLG 1560  
Db 2526 DVNRLSSKAVNHRSVWKKLLEDTEPIDTTIMAKNEFCVQPEKGRKPARLIIVFPDLG 2585  
QY 1561 VVCEKMALYDVVSTLPQAVMGSSYGFQYSPQGRVEFLVNAWKAKCPMGFYAYTRCFDS 1620  
Db 2586 VVCEKMALYDVVSTLPQAVMGSSYGFQYSPQGRVEFLVNAWKAKCPMGFYAYTRCFDS 2645  
QY 1621 TVTENDIRVESIYOCCLDAPARQAIRSLTERLYTGGPLTNSKQNCQYRCRASGVLT 1680  
Db 2646 TVTENDIRVESIYOCCLDAPARQAIRSLTERLYTGGPLTNSKQNCQYRCRASGVLT 2705  
QY 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVCESAGTQEDASLRAFTAMTRY 1740  
Db 2706 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVCESAGTQEDASLRAFTAMTRY 2765  
QY 1741 SAPPDGPPEYDLELITSCSNVSVAHDAKGRVYLLTRDPTTLARAANETARTFPVN 1800  
Db 2766 SAPPDGPPEYDLELITSCSNVSVAHDAKGRVYLLTRDPTTLARAANETARTFPVN 2825  
QY 1801 SWLGNIMVAPTLWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIQRL 1860  
Db 2826 SWLGNIMVAPTLWARMILMTHFFSILLAEQLEKALDCQIYGACYSIEPLDLPQIIQRL 2885  
QY 1861 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRWHRARSVRAILLSQGGRAATCKGYLF 1920  
Db 2886 HGLSAFSLHSYSGEINRVASCLRKLGVPPLRWHRARSVRAILLSQGGRAATCKGYLF 2945  
QY 1921 NNAVTKLTLTPAASOLDLSNFWFVAGYSGGDIYHLSRARPFRMFWCLLILSVGVGY 1980  
Db 2946 NNAVTKLTLTPAASOLDLSNFWFVAGYSGGDIYHLSRARPFRMFWCLLILSVGVGY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

RESULT 7

Q9DTE7  
ID Q9DTE7 PRELIMINARY; PRT; 3010 AA.  
AC Q9DTE7;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Serum;  
RA Takahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,  
Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,  
Mishiro S.;  
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients  
with hepatocellular carcinoma: the 'progression score' revisited.";  
RL Hepatol. Res. 20:161-171(2001).  
DR EMBL; AB049090; BAB18803.1; -;  
DR PIR; A61196; A61196.  
DR PIR; P00246; P00246.  
DR PIR; P00252; P00252.  
DR PIR; P00253; P00253.  
DR PIR; P00254; P00254.  
DR PIR; P00255; P00255.  
DR PIR; P00256; P00256.  
DR HSP; P26663; LUXP.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0019028; C:viral capsid; IEA.  
DR GO; GO:0019031; C:viral envelope; IEA.  
DR GO; GO:0005524; F:ATP binding; IEA.  
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO; GO:0003723; F:RNA binding; IEA.  
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO; GO:0005198; F:structural molecule activity; IEA.  
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.  
DR GO; GO:0006350; P:transcription; IEA.  
DR GO; GO:0001907; F:viral genome replication; IEA.  
DR GO; GO:0019087; F:viral transformation; IEA.  
DR InterPro; IPR000345; CytC\_heme\_BS.  
DR InterPro; IPR001410; DEAD.  
DR InterPro; IPR002522; HCV capsid.  
DR InterPro; IPR002521; HCV core.  
DR InterPro; IPR002519; HCV env.  
DR InterPro; IPR002531; HCV NS1.  
DR InterPro; IPR000745; HCV NS4a.  
DR InterPro; IPR001490; HCV NS4b.  
DR InterPro; IPR002868; HCV NS5a.  
DR InterPro; IPR002166; HCV RdRP.  
DR InterPro; IPR001650; Helicase\_C.  
DR InterPro; IPR004109; Peptidase\_S29.  
DR InterPro; IPR009003; Pept\_Ser\_Cys.  
DR InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR Pfam; PF01543; HCV capsid; 1.  
DR Pfam; PF01542; HCV core; 1.  
DR Pfam; PF01539; HCV env; 1.  
DR Pfam; PF01560; HCV NS1; 1.  
DR Pfam; PF01538; HCV NS2; 1.  
DR Pfam; PF02907; HCV NS3; 1.  
DR Pfam; PF01006; HCV NS4a; 1.  
DR Pfam; PF01001; HCV NS4b; 1.  
DR Pfam; PF01506; HCV NS5a; 1.  
DR Pfam; PF00271; Helicase\_C; 1.  
DR Pfam; PF00998; Viral RdRp; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN\_1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
Polyprotein; Transmembrane.

SQ	SEQUENCE	3010 AA; 326838 MW; 58E3BD4140B588 CRC64;
	Query Match	97.6%; Score 10218; DB 2; Length 3010;
	Best Local Similarity	96.7%; Pred. No. 0;
	Matches 1920; Conservative	39; Mismatches 26; Indels 0; Gaps 0;
QY	1	MAPITAYSOQTRGLGCIITSLGRDRNOVEGEVONVSTATQSEFLATCVNGVCWTVYHGA 60
DB	1026	LAPITAYSOQTRGLGCIITSLGRDRNOVEGEVONVSTATQSEFLATCVNGVCWTVYHGA 1085
QY	61	GSKTLAGKPGPIITQMYINVQDQVLGVWQAPPGARSLTPTCTCGSSDLYLVTRHADVIPVRRR 120
DB	1086	GSKTLAGKPGPIITQMYINVQDQVLGVWQAPPGARSLTPTCTCGSSDLYLVTRHADVIPVRRR 1145
QY	121	GDSRGSLLSPRPVSYLKGSSGGPILLCPSPGHAVGIFPRAAVCTRGVAKAVDFVVPVSMETTM 180
DB	1146	GDSRGSLLSPRPVSYLKGSSGGPILLCPSPGHAVGIFPRAAVCTRGVAKAVDFVVPVSMETTM 1205
QY	181	RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAQAQYKVLVNPVSAATLFGF 240
DB	1206	RSPVFTDNSSPPAVPQTFQVAHLHAPTSGSKSTKVPAAQAQYKVLVNPVSAATLFGF 1265
QY	241	AYMSKAHGIDENIRTVRTITTTGAPITYSYVYKFLADGGCGGAYDIIIIIDECBSTDSTT 300
DB	1266	AYMSKAHGIDENIRTVRTITTTGAPITYSYVYKFLADGGCGGAYDIIIIIDECBSTDSTT 1325
QY	301	ILGIGTVLDOAETAGARLVVATATPPGSVTVPHNPTEEVALSSTGBIPYGAIPETI 360
DB	1326	ILGIGTVLDOAETAGARLVVATATPPGSVTVPHNPTEEVALSSTGBIPYGAIPETI 1385
QY	361	KGRHLIFCHSKKCDLAAKSLGLGNVAYVYRGLDVSIVPTSGDVIVVATDALMTGFT 420
DB	1386	KGRHLIFCHSKKCDLAAKSLGLGNVAYVYRGLDVSIVPTSGDVIVVATDALMTGFT 1445
QY	421	GFDSVIDNCNTCVTQTVDFSLDPTFTTETTVPODAVRSQRGRGTGRGMGIYRFVTPG 480
DB	1446	GFDSVIDNCNTCVTQTVDFSLDPTFTTETTVPODAVRSQRGRGTGRGMGIYRFVTPG 1505
QY	481	ERSGMPDSSVLCEDYDAGAWBELTAPETSRLRAYLNTPGIPVQDHLFEFVESVPTGL 540
DB	1506	ERSGMPDSSVLCEDYDAGAWBELTAPETSRLRAYLNTPGIPVQDHLFEFVESVPTGL 1565
QY	541	THIDAFSLQTKAGDNFPVLYVQATVCARAQAPPSWDMKCLRLKPTLHGPTPL 600
DB	1566	THIDAFSLQTKAGDNFPVLYVQATVCARAQAPPSWDMKCLRLKPTLHGPTPL 1625
QY	601	YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVIAGR 660
DB	1626	YRLGAVQNEVTTTHPTIKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSVVIAGR 1685
QY	661	IILSGKPAITPDREVLYREFDEMEECASHLPYIEQGMQLAEQKQKALGLQTATKQAEA 720
DB	1686	IILSGKPAITPDREVLYREFDEMEECASHLPYIEQGMQLAEQKQKALGLQTATKQAEA 1745
QY	721	AAPVSKWTLLEAFWAKHWNFTSGIYQYLAGLSTLPGNPAIASLMAFTASITSPLTQ 780
DB	1746	AAPVSKWTLLEAFWAKHWNFTSGIYQYLAGLSTLPGNPAIASLMAFTASITSPLTQ 1805
QY	781	TLLFNILGWVAAQLAPPSAASAFVAGIAGAAVGSIGLKVLDVILAGYAGVAGALVA 840
DB	1806	TLLFNILGWVAAQLAPPSAASAFVAGIAGAAVGSIGLKVLDVILAGYAGVAGALVA 1865
QY	841	FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQWNNRLIAFASR 900
DB	1866	FKVMSGEMPTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVQWNNRLIAFASR 1925
QY	901	GNHVSPTHYVPESDAAARVQIILSSLTITOLLKELHOMINEDCSTPCSGSLRDVNDWIC 960
DB	1926	GNHVSPTHYVPESDAAARVQIILSSLTITOLLKELHOMINEDCSTPCSGSLRDVNDWIC 1985
QY	961	TVLTDFKTLQSKLLPRLPGVPFFSCQYKGYVWRGDGIMQTTCPGCAQITGHVXGSMR 1020
DB	1986	TVLTDFKTLQSKLLPRLPGVPFFSCQYKGYVWRGDGIMQTTCPGCAQITGHVXGSMR 2045

QY	1021	IVGPTCSNTWHTGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEVTRVGDPHYVTGM 1080
DB	2046	IVGPTCSNTWHTGTFPINAYTTGCTPSPAPNYSRALWRVAABEYVEITRVGDPHYVTGM 2105
QY	1081	TTNVKPCQOVPAPEFTEVDGVLHRYAPACKPLAREEYTELVLGNOYLVGSOLPCEPE 1140
DB	2106	TTNVKPCQOVPAPEFTELDGVLHRYAPACKPLAREEYTFQGLNQYVVGSLPCEPE 2165
QY	1141	PDVAVLTSMLTDSHITAEAKERLARGSPSPSLASSASQSLAPSLKATCTTRHDSDDAD 1200
DB	2166	PDVAVLTSMLTDSHITAEAKERLARGSPSPSLASSASQSLAPSLKATCTTRHDSDDAD 2225
QY	1201	LIEANLLWRQEMGNITRVSEKENVILDSFEPLOABEDREVSVPABILRRSKFFPRAM 1260
DB	2226	LIEANLLWRQEMGNITRVSEKENVILDSFEPLOABEDREVSVAEILRRSKFFPAL 2285
QY	1261	PIWARPNDYNPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA 1320
DB	2286	PIWARPNDYNPLLESWKDPDYVPPVHGCPLPAKAPPIPPRRKRTVVLSESTVSSALA 2345
QY	1321	ELATKTFGSSSESAVDSGTATASPDQPSDDGDAGSDVESYSSMPPLEGEPCDPLDSGWS 1380
DB	2346	ELATKTFGSSSESAVDSGTATASPDQPSDDGDAGSDVESYSSMPPLEGEPCDPLDSGWS 2405
QY	1381	STVSEBSEASDVVCCSMSYTWGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS 1440
DB	2406	STVSEBSEASDVVCCSMSYTWGALITPCAABETKLPINALSNSLLRHHNLVYATTSRAS 2465
QY	1441	LROKVTTFDRLQVLDHRYDLVKEMKAKASTVAKLLSVEEACKLTPPHSARSFGYGAK 1500
DB	2466	LROKVTTFDRLQVLDHRYDLVKEMKAKASTVAKLLSVEEACKLTPPHSARSFGYGAK 2525
QY	1501	DVNLSKAVNHHSVWVKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560
DB	2526	DVNLSKAVNHHSVWVKOLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 2585
QY	1561	VRVCEKMALYDVVSTLTPQAVMGSSYGFQVSPGQVRFVFNNAWAKKCPMGFAYDTRCFDS 1620
DB	2586	VRVCEKMALYDVVSTLTPQAVMGSSYGFQVSPGQVRFVFNNAWAKKCPMGFAYDTRCFDS 2645
QY	1621	TVTENDIRVEESYQCCDLAPEARQALRSILTERLYIGGPLTNSKNGONCGYRRCRAGSVLT 1680
DB	2646	TVTENDIRVEESYQCCDLAPEARQALRSILTERLYIGGPLTNSKNGONCGYRRCRAGSVLT 2705
QY	1681	TSCGNTLTCYKAAAACRAAKLQDCIMLVGDDLWICESAGTQDEASLRATEAMTRY 1740
DB	2706	TSCGNTLTCYKAAAACRAAKLQDCIMLVGDDLWICESAGTQDEASLRATEAMTRY 2765
QY	1741	SAPPGPDPPEYDLELITSCSSNVVAHDASGRVYVYLTDRDPTPLARAANETARHTPVN 1800
DB	2766	SAPPGPDPPEYDLELITSCSSNVVAHDASGRVYVYLTDRDPTPLARAANETARHTPVN 2825
QY	1801	SWLGNIMVAPTILWARMILMTHFFSILLAOEQLKALDCOYGCACYSEIPLDLPQIIRL 1860
DB	2826	SWLGNIMVAPTILWARMILMTHFFSILLAOEQLKALDCOYGCACYSEIPLDLPQIIRL 2885
QY	1861	HGLSAPLSHSYSPGEINRVASCLRLKGLVPPPLRVHRARSVRARLLSQGGRATCGKLYF 1920
DB	2886	HGLSAPLSHSYSPGEINRVASCLRLKGLVPPPLRVHRARSVRARLLSQGGRATCGKLYF 2945
QY	1921	NWAVRKLKLTPTIPAAASQDLSSNFVAGYGGDIYHSLSRARPRPFWMLLLLSVGVIY 1980
DB	2946	NWAVRKLKLTPTIPAAASQDLSSNFVAGYGGDIYHSLSRARPRPFWMLLLLSVGVIY 3005
QY	1981	LLPNR 1985
DB	3006	LLPNR 3010

RESULT 8  
Q7T4V8  
ID Q7T4V8 PRELIMINARY; PRT; 1984 AA.



Db 1381 TVSEASEDVVCCSMYSYTWGALITPCAAEESKLPINALSNSLLRHHNMVYATTSSASQ 1440  
QY 1442 RQKVTFRLOVLDHDDYDLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKD 1501  
Db 1441 RQKVTFRLOVLDHDDYDLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAKD 1500  
QY 1502 VRNLSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVCVQPKGGRKPARLIVFPDLGV 1561  
Db 1501 VRNLSKAVNHRSVWKKLLEDTETPIDTTIMAKNEVCVQPKGGRKPARLIVFPDLGV 1560  
QY 1562 RVCEKMALYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDST 1621  
Db 1561 RVCEKMALYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDST 1620  
QY 1622 VTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLETNSKQNCYRCRASGVLT 1681  
Db 1621 VTENDIRVEESIYQCCDLAPARQAIRSLTERLYIGGPLETNSKQNCYRCRASGVLT 1680  
QY 1682 SCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY 1741  
Db 1681 SCGNLTLCYLKASACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTAMTRY 1740  
QY 1742 APPGDPKPEYDLLEITSCSNVSVAHDAAGKRVYLLTRDPTTFLARAAMETARHTPVNS 1801  
Db 1741 APPGDPKPEYDLLEITSCSNVSVAHDAAGKRVYLLTRDPTTFLARAAMETARHTPVNS 1800  
QY 1802 WLGNIMVAPTLWARMLTWHFFSILLAOEQLKALDCQIYGACYSIEPDLPLQIIORLH 1861  
Db 1801 WLGNIMVAPTLWARMLTWHFFSILLAOEQLKALDCQIYGACYSIEPDLPLQIIORLH 1860  
QY 1862 GLSAFSLHSYSPGEINRVASCLRLKGLVPPPLRVWHRARSVRARLLSOGGRAATCGKYLEN 1921  
Db 1861 GLSAFSLHSYSPGEINRVASCLRLKGLVPPPLRVWHRARSVRARLLSOGGRAATCGKYLEN 1920  
QY 1922 WAVRTKLKLTPIPAASQDLSSWVAGYSGGDIYVHLSRARPRFWMWCLLLSVGVGIYL 1981  
Db 1921 WAVRTKLKLTPIPAASRLDLSGWVAGYSGGDIYVHLSRARPRFWMWCLLLSVGVGIYL 1980  
QY 1982 LPNR 1985  
Db 1981 LPNR 1984

RESULT 9  
P90194 PRELIMINARY; PRT; 3010 AA.  
AC P90194;  
DT 01-MAY-1997 (TrEMBLrel. 03, Created)  
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)  
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]\_SEQUENCE FROM N.A.  
RP STRAIN=HCV-1b;  
RC MEDLINE=95340824; PubMed=7542279;  
RX Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T., Yamamoto C., Izumi N., Marumo F., Sato C.;  
RA "Comparison of full-length sequences of interferon-sensitive and resistant hepatitis C virus 1b";  
RT J. Clin. Invest. 96:224-230 (1995).  
RL [2]  
RN SEQUENCE FROM N.A.  
RP STRAIN=HCV-1b;  
RC Enomoto N.;  
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.  
DR EMBL; D50481; BAA09072.1; -.  
DR FIR; A61196; A61196.  
DR FIR; PQ0246; PQ0246.  
DR FIR; PQ0254; PQ0254.

PIR; PS0329; PS0329.  
HSP; P26663; LUXP.  
GO; GO:0016021; C: integral to membrane; IEA.  
GO; GO:0019028; C: viral capsid; IEA.  
GO; GO:0019031; C: viral envelope; IEA.  
GO; GO:0005524; F: ATP binding; IEA.  
GO; GO:0008026; F: ATP-dependent helicase activity; IEA.  
GO; GO:0003723; F: RNA binding; IEA.  
GO; GO:0008236; F: RNA-directed RNA polymerase activity; IEA.  
GO; GO:0005198; F: serine-type peptidase activity; IEA.  
GO; GO:0006508; P: proteolysis and peptidolysis; IEA.  
GO; GO:0006350; P: transcription; IEA.  
GO; GO:0019079; P: viral genome replication; IEA.  
GO; GO:0019087; P: viral transformation; IEA.  
InterPro; IPR000345; CytC\_heme\_BS.  
InterPro; IPR001410; DEAD.  
InterPro; IPR002522; HCV\_capsid.  
InterPro; IPR002519; HCV\_core.  
InterPro; IPR002531; HCV\_NS1.  
InterPro; IPR000745; HCV\_NS4a.  
InterPro; IPR001490; HCV\_NS4b.  
InterPro; IPR002868; HCV\_NS5a.  
InterPro; IPR002166; HCV\_RdRP.  
InterPro; IPR001650; Helicase\_C.  
InterPro; IPR004109; Peptidase\_S29.  
InterPro; IPR009003; Peptidase\_S29.  
InterPro; IPR002518; Pept\_U39\_HCV\_NS2.  
InterPro; IPR007095; RNA\_pol\_DS\_PS.  
InterPro; IPR007094; RNA\_pol\_PSVir.  
Pfam; PF01543; HCV\_capsid; 1.  
Pfam; PF01542; HCV\_core; 1.  
Pfam; PF01539; HCV\_env; 1.  
Pfam; PF01560; HCV\_NS1; 1.  
Pfam; PF01538; HCV\_NS2; 1.  
Pfam; PF02907; HCV\_NS3; 1.  
Pfam; PF01006; HCV\_NS4a; 1.  
Pfam; PF01001; HCV\_NS4b; 1.  
Pfam; PF01506; HCV\_NS5a; 1.  
Pfam; PF00271; Helicase\_C; 1.  
Pfam; PF00998; Viral\_RdRP; 1.  
SMART; SM00487; DEXDC; 1.  
PROSITE; PS00190; CYTOCHROME\_C; UNKNOWN 1.  
Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
Polyprotein; Transmembrane.  
CHAIN 810 1026 NS2.  
CHAIN 1027 1657 NS3.  
CHAIN 1658 1711 NS4a.  
CHAIN 1 191 core protein.  
CHAIN 1712 1972 NS4b.  
CHAIN 1973 2419 NS5a.  
CHAIN 2420 3010 NS5b.  
CHAIN 192 383 E1.  
CHAIN 384 809 E2.  
SEQUENCE 3010 AA; 326816 MW; 98D5C2A2D47FD011 CRC64;  
Query Match 97.5%; Score 10199; DB 2; Length 3010;  
Best Local Similarity 96.7%; Pred. No. 0;  
Matches 1920; Conservative 34; Mismatches 31; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLLCIITSLTGRDKNQVEGEVQVWSTATQSFATCNGVWYVHGA 60  
1026 LAPITAYSQOTRGLLCIITSLTGRDKNQVEGEVQVWSTATQSFATCNGVWYVHGA 1085  
Db 61 GSKTLGPKGPITQMTYTNVDQLVGNQAPPGARSLTPTCGSSDLYLVTRHADVIPVRR 120  
1086 GSKTLGPKGPITQMTYTNVDQLVGNQAPPGARSLTPTCGSSDLYLVTRHADVIPVRR 1145  
QY 121 GDSRGLSLSPRVSYLKGSSGGLPLCPSPHAGVIFRAAVCTRGVAKAVDFVPSMETTM 180  
1146 GDSRGLSLSPRVSYLKGSSGGLPLCPSPHAGVIFRAAVCTRGVAKAVDFVPSMETTM 1205

191 RSPVPTDSSPPAVPQTTOVAHLHAPTGSKSTKVPAAAYAGYKVVLNPSVAATI.GFG 240  
1206 RSPVPTDSSPPAVPQTTOVAHLHAPTGSKSNKVPVEYAAQGYKVVLNPSVAATL.GFG 1265  
241 AYMSKAHGIIDPNIRTVGRTITTGAPITTYSTYKFLADGGCSGAYDIIICDCHSTDSST 300  
1266 AYMSKAHGVDPNIRTVGRTITTGAPITTYSTYKFLADGGCSGAYDIIICDCHSTDSST 1325  
301 ILIGTIVLDOAQTAGARLVLTATATPPGSVTVPHENIEVALSSTGEIPFYGKA.PIETI 360  
1326 ILIGTIVLDOAQTAGARLVLTATATPPGSVTVPHENIEVALSSTGEIPFYGKA.PIETI 1385  
361 KGRHHLIFCHSKKKCDELAAKLSGLGNAYVYRGLDIVSVIPTSGDVIVWATDALMTGFT 420  
1386 KGRHHLIFCHSKKKCDELAAKLSGLGNAYVYRGLDIVSVIPTSGDVIVWATDALMTGFT 1445  
421 GPDFSDVIDCNTCTVOTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGRGMGIYRFVTPG 480  
1446 GPDFSDVIDCNTCTVOTVDFSLDPTFTIETTTVPQDAVSRSQRRGRTGRGMGIYRFVTPG 1505  
481 ERPSGMDSSVLCCEYDAGCAYVELTPAETSRLRAYLNTPGLPVQDHLFEWESVFTGL 540  
1506 ERPSGMDSSVLCCEYDAGCAYVELTPAETSRLRAYLNTPGLPVQDHLFEWESVFTGL 1565  
541 THIDAHFLSOTQAGDNFPYLYAYQATVCARAQAPPSWDQWKCLIRLKPTLHGPTLL 600  
1566 THIDAHFLSOTQAGDNFPYLYAYQATVCARAQAPPSWDQWKCLIRLKPTLHGPTLL 1625  
601 YRLGAVQNEVTTTHPTIKYIMACMSADLEVTSTWLVGVLAALAAAYCLITTSVVIVGR 660  
1626 YRLGAVQNEVTTTHPTIKYIMACMSADLEVTSTWLVGVLAALAAAYCLITTSVVIVGR 1685  
661 IILSGKPAIIPDREVLRYREFDEMEECASHLYPIEQMQLAEOFKQKALGLLOTATQKAEA 720  
1686 IILSGKPAIIPDREVLRYREFDEMEECASHLYPIEQMQLAEOFKQKALGLLOTATQKAEA 1745  
721 AAPVSKWRTLEAFWAKHWNFIISGIVLAGLSTLPGNPATASLMAFTASITSP.LTTOH 780  
1746 AAPVSKWRTLEAFWAKHWNFIISGIVLAGLSTLPGNPATASLMAFTASITSP.LTTOH 1805  
781 TLLFNILGCVAAQLAPPASAASAFVAGIAGAAVGSIGLGKVLVDILAGYAGVAGALVA 840  
1806 TLLFNILGCVAAQLAPPASAASAFVAGIAGAAVGSIGLGKVLVDILAGYAGVAGALVA 1865  
841 FKVMGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGECAVQWNNRLIAPASR 900  
1866 FKVMGEMPTEDLVNLLPAILSPGALVGVCAAILRRHVGPGECAVQWNNRLIAPASR 1925  
901 GNVHSPTHVYVESDAARVTOILSLTITQLKRLHWINEDCSTPCSGSWLRD.VMDWIC 960  
1926 GNVHSPTHVYVESDAARVTOILSLTITQLKRLHWINEDCSTPCSGSWLRD.VMDWIC 1985  
961 TVLTDFKTLQSKLLPRLPGVPFPFCQYGYKVMRGDGMOTTCPCGAQITGHVKXGSMR 1020  
1986 TVLTDFKTLQSKLLPRLPGVPFPFCQYGYKVMRGDGMOTTCPCGAQITGHVKXGSMR 2045  
1021 IVGPRTCNSNTHGCTFPINAYTTGCTPSPAPNYSRALRWVAABEYVEVTRGVFHYVTGM 1080  
2046 IVGPRTCNSNTHGCTFPINAYTTGCTPSPAPNYSRALRWVAABEYVEVTRGVFHYVTGM 2105  
1081 TTDNVKCPQVPAPEPFEVDGVLHRYAPACKPLLRREYVTVLVLGNQYLVGSQ.LPCEPE 1140  
2106 TTDNVKCPQVPAPEPFEVDGVLHRYAPACKPLLRREYVTVLVLGNQYLVGSQ.LPCEPE 2165  
1141 PDVAVLTSMLTDFSHITAEATKRLARGSPSPSLASSASQLSAPSLKATCTTHR.DSPAD 1200  
2166 PDVAVLTSMLTDFSHITAEATKRLARGSPSPSLASSASQLSAPSLKATCTTHR.DSPAD 2225  
1201 LIEANLLWRQMGNNITRVSENNKVILDSFEPLQAEEDEREVSVPAAEILRRSKKPPRAM 1260  
2226 LIEANLLWRQMGNNITRVSENNKVILDSFEPLQAEEDEREVSVPAAEILRRSKKPPRAM 2285  
1261 PIWARPDPYNPLESWKDPDVPVPPVHVCPLPPAKAPPIPPRRKRTVWLSESTVSSALA 1320

2286 PIWARPDPYNPLESWKDPDVPVPPVHVCPLPPAKAPPIPPRRKRTVWLSESTVSSALA 2345  
1321 ELATKTFGSSSAVDSCGTATASPDQDDGAGDVESYSMPLEGEPPDLS.DSGSW 1380  
2346 ELATKTFGSSSAVDSCGTATASPDQDDGAGDVESYSMPLEGEPPDLS.DSGSW 2405  
1381 STVSEAESEDDVCCSWSYTWTCALITPCAABETKLPINALSNLLRHHNVVATT.SRSAS 1440  
2406 STVSEAESEDDVCCSWSYTWTCALITPCAABESKLPINALSNLLRHHNVVATT.SRSAS 2465  
1441 LRQKXVTEDRLQVLDHDDHYRDVLKEMKAKASTVKAKLSVEEACKLTPPHSARSKFYGAK 1500  
2466 LRQKXVTEDRLQVLDHDDHYRDVLKEMKAKASTVKAKLSVEEACKLTPPHSARSKFYGAK 2525  
1501 DVNRNLSKAVNHIRSVWKOLLEDTETPTDITTIMAKNEVFCVQPEKGRKPAELI.VFPDLG 1560  
2526 DVNRNLSKAVNHIRSVWKOLLEDTETPTDITTIMAKNEVFCVQPEKGRKPAELI.VFPDLG 2585  
1561 VRVCEKMAIYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDS 1620  
2586 VRVCEKMAIYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDS 2645  
1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCRAGSVLT 1680  
2646 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGPLTNSKGQNCYRRCRAGSVLT 2705  
1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTOEDEASIRAFTEAMTAY 1740  
2706 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVVICESAGTOEDEASIRAFTEAMTAY 2765  
1741 SAPPGDPKPEYDLELITSCSSNVSVAHDSAGKRVVYLTRDPTPLARA.AWETARTHPVN 1800  
2766 SAPPGDPKPEYDLELITSCSSNVSVAHDSAGKRVVYLTRDPTPLARA.AWETARTHPVN 2825  
1801 SWLGNIIIMVAPTILWARMILMTHFFSILLAOEQLEKALDQCIYGCACYSIEPLDLP.QIQR 1860  
2826 SWLGNIIIMVAPTILWARMILMTHFFSILLAOEQLEKALDQCIYGCACYSIEPLDLP.QIQR 2885  
1861 HGLSAPLSHSYSGEINRVASCLIKLGVPLVWRHRSVRARLLSOGGRAATCGKLYF 1920  
2886 HGLSAPLSHSYSGEINRVASCLIKLGVPLVWRHRSVRARLLSOGGRAATCGKLYF 2945  
1921 NNAVTRKLTLPAPASQDLSSWFSVAGYSGGDIHSLSRARPRFWFWCCLLLSVGVGIY 1980  
2946 NNAVTRKLTLPAPASQDLSSWFSVAGYSGGDIHSLSRARPRFWFWCCLLLSVGVGIY 3005  
1981 LLPNR 1985  
3006 LLPNR 3010

## RESULT 10

Q99AU2 PRELIMINARY; PRT: 3010 AA.  
ID Q99AU2  
AC Q99AU2;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)  
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus type 1b.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus; Hepatitis C virus type 1.  
OX NCBI\_TaxID=31647;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=chimera of HCV-BK;  
RX MEDLINE=21534507; PubMed=11677216;  
RA Thomson M., Nascimbeni M., Gonzales S., Murthy K.K., Rehmann B.,  
RT Liang T.J.;  
RT "Emergence of a distinct pattern of viral mutations in chimpanzees  
infectd with a homogeneous inoculum of hepatitis C virus.";  
RL Gastroenterology 121:1226-1233(2001).





1381 STVSEASEDVVCCSMYSWTGALITPCAAEBETKLPINALSNLLRHNLVYATTSRSAS 1440  
2406 STVSEASEDVVCCSMYSWTGALITPCAAEBESKLPINALSNLLRHNLVYATTSRSAG 2465  
1441 LQKQVTFDLQVLDHVDYDLVLEKMAKASTYKAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
2466 LQKQVTFDLQVLDHVDYDLVLEKMAKASTYKAKLLSVEEACKLTPPHSARSKFGYGAK 2525  
1501 DYRNLSKAVNHRSWKDLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIIVPDLG 1560  
2526 DYRNLSKAVNHRSWKDLLEDTETPIDTTIMAKNEVFCVQPEKGRKPARLIIVPDLG 2585  
1561 VVACEKMALYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNAWKAKCPEMFAYDTRCFDS 1620  
2586 VVACEKMALYDVVSTLPQAVMGSSYGFQYSPQORVEFLVNTWKSXKPMGFSDYTRCFDS 2645  
1621 TVTENDIRVEESIYQCCDLAPAPARQAIRSLTERLYIGGELTNSKGONCGYRRCRASGVLT 1680  
2646 TVTENDIRVEESIYQCCDLAPAPARQAIRSLTERLYIGGELTNSKGONCGYRRCRASGVLT 2705  
1681 TSCGNTLTCLYKAAACRAAKLQDCTMLVCGDGLVVICESAGTOEDASLRAFTAMTRY 1740  
2706 TSCGNTLTCLYKAAACRAAKLQDCTMLVNGDDLWVICESAGTOEDAAASLRAFTAMTRY 2765  
1741 SAPPDGPPEVDLELITSCSNVSVAHDAAGKRVYVLTTRDPTTPLARAAMETARHTPVN 1800  
2766 SAPPDGPPEVDLELITSCSNVSVAHDAAGKRVYVLTTRDPTTPLARAAMETARHTPVN 2825  
1801 SWLGNIMYAPTLMWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPQIQL 1860  
2826 SWLGNIMYAPTLMWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPQIQL 2885  
1861 HGLSAFSLHSYSPGEINRVASCLKLGVPPLVWRHRSVBARLLSQGGRATCGKLYF 1920  
2886 HGLSAFSLHSYSPGEINRVASCLKLGVPPLVWRHRSVBARLLSQGGRATCGKLYF 2945  
1921 NNAVATKILKLTIPAAQDLSSWVAGYSGGDIYHLSRARPWFMMCLLLLSVGVGTY 1980  
2946 NNAVATKILKLTIPAAQDLSSWVAGYSGGDIYHLSRARPWFMMCLLLLSVGVGTY 3005  
1981 LLPNR 1985  
3006 LLPNR 3010

RESULT 11

Q9DTE9  
ID Q9DTE9 PRELIMINARY; PRT: 3010 AA.  
AC Q9DTE9;  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
DE Polyprotein.  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;  
RN [1]\_TaxID=11103;  
RP SEQUENCE FROM N.A.  
RC TISSUE=Serum;  
RA Hatahashi K., Iwata K., Matsumoto M., Matsumoto H., Nakao K.,  
RA Hatahara T., Ohta Y., Kanai K., Maruo H., Baba K., Hijikata M.,  
RA Mishiro S.;  
RT "Hepatitis C virus (HCV) genotype 1b sequences from fifteen patients  
RT with hepatocellular carcinoma: the 'progression score' revisited."  
RL Hepatol. Res. 20:161-171(2001).  
DR EMBL; AB049086; BAB18801.1; -.  
DR PIR; A61196; A61196.  
DR PIR; PQ0246; PQ0246.  
DR PIR; PQ0804; PQ0804.  
DR PIR; PS0329; PS0329.  
DR HSSP; P26663; LJXP.  
GO; GO:0016021; C:integral to membrane; IEA.

DR GO:0019028; C:viral capsid; IEA.  
DR GO:0019031; C:viral envelope; IEA.  
DR GO:0005524; F:ATP binding; IEA.  
DR GO:0008026; F:ATP-dependent helicase activity; IEA.  
DR GO:0003723; F:RNA binding; IEA.  
DR GO:0003968; F:RNA-directed RNA polymerase activity; IEA.  
DR GO:0008236; F:serine-type peptidase activity; IEA.  
DR GO:0005198; F:structural molecule activity; IEA.  
DR GO:000508; F:proteolysis and peptidolysis; IEA.  
DR GO:0006350; F:transcription; IEA.  
DR GO:0019079; F:viral genome replication; IEA.  
DR GO:0019087; F:viral transformation; IEA.  
DR InterPro: IPR00345; CytC\_heme\_BS.  
DR InterPro: IPR001410; DEAD\_capsid.  
DR InterPro: IPR002522; HCV\_capsid.  
DR InterPro: IPR002521; HCV\_core.  
DR InterPro: IPR002519; HCV env.  
DR InterPro: IPR002531; HCV NS1.  
DR InterPro: IPR00745; HCV NS4a.  
DR InterPro: IPR001490; HCV NS4b.  
DR InterPro: IPR002868; HCV NS5a.  
DR InterPro: IPR002166; HCV RdRp.  
DR InterPro: IPR001650; Helicase\_C.  
DR InterPro: IPR004109; Peptidase\_S29.  
DR InterPro: IPR009003; Pept\_Ser\_Cys.  
DR InterPro: IPR002518; Pept\_U39\_HCV\_NS2.  
DR InterPro: IPR007095; RNA\_pol\_BS\_PS.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR Pfam: PF01543; HCV capsid; 1.  
DR Pfam: PF01542; HCV core; 1.  
DR Pfam: PF01539; HCV env; 1.  
DR Pfam: PF01560; HCV NS1; 1.  
DR Pfam: PF01538; HCV NS2; 1.  
DR Pfam: PF02907; HCV NS3; 1.  
DR Pfam: PF01006; HCV NS4a; 1.  
DR Pfam: PF01001; HCV NS4b; 1.  
DR Pfam: PF01506; HCV NS5a; 1.  
DR Pfam: PF00271; Helicase\_C; 1.  
DR Pfam: PF00998; Viral\_RdRp; 1.  
DR SMART; SM00487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
SQ SEQUENCE 3010 AA; 326780 MW; 668CFFEA5FEC3658 CRC64;  
Query Match 97.3%; Score 10187; DB 2; Length 3010;  
Best Local Similarity 96.5%; Pred.No. 0;  
Matches 1915; Conservative 41; Mismatches 29; Indels 0; Gaps 0;  
QY 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVWSTATQSLATCVNGVCTVYHGA 60  
DB 1026 LAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVWSTATQSLATCVNGVCTVYHGA 1085  
QY 61 GSKTLIAGKGPITQMYTNVDQLVQWQAPPGARSLTPTCGSSDLVLTTRHADVIPVRR 120  
DB 1086 GSKTLIAGKGPITQMYTNVDQLVQWQAPPGARSLTPTCGSSDLVLTTRHADVIPVRR 1145  
QY 121 GDSRGLSPRPVSYLKGSSGGLCPGSHAVGIFFRAAVCTRGAKAVDFVVESEMTTM 180  
DB 1146 GDSRGLSPRPVSYLKGSSGGLCPGSHAVGIFFRAAVCTRGAKAVDFVVESEMTTM 1205  
QY 181 RSPVFTDNSSPPAVPOTQVAHLHAPTQSGKSTKVPAAAYAAQYKVLVNPVVAATLFGP 240  
DB 1206 RSPVFTDNSSPPAVPOTQVAHLHAPTQSGKSTKVPAAAYAAQYKVLVNPVVAATLFGP 1265  
QY 241 AYMSKAHGIDNPRTGVRTITTTGAPITYSTYTGKFLADGCGSGAYDIIICDECHSDSTT 300  
DB 1266 AYMSKAHGIDNPRTGVRTITTTGAPITYSTYTGKFLADGCGSGAYDIIICDECHSDSTT 1325  
QY 301 ILGIGTVDLQDAETAGARLVVLATATPPGSVTVPHNIEVALSSTGEIIFYGKAIPETI 360  
DB 1326 ILGIGTVDLQDAETAGARLVVLATATPPGSVTVPHNIEVALSSTGEIIFYGKAIPETI 1385

```
QY 361 KGRHLLFCHSKKKDELAALKSLGLGNAYAYRGLDVSIVPTSGDIVVATDALMTGFT 420
Db 1386 KGRHLLFCHSKKKDELAALKSLGLGNAYAYRGLDVSIVPTSGDIVVATDALMTGFT 1445
QY 421 GDFDSVIDCNTCVTQTQVDFSLDPTFTTETTPQDAVSRORRGTGRGMGIYRFVTPG 480
Db 1446 GDFDSVIDCNTCVTQTQVDFSLDPTFTTETTPQDAVSRORRGTGRGMGIYRFVTPG 1505
QY 481 ERPSGMDSSVLCBYDAGCAWYELTPTAETSVRLRAYINTPGLPVCQDHLFEWESVFTGL 540
Db 1506 ERPSGMDSSVLCBYDAGCAWYELTPTAETSVRLRAYINTPGLPVCQDHLFEWESVFTGL 1565
QY 541 THIDAHFLSOTKQAGDNFYLVAQATVCARAQAPPSWQMKCLIRLKPTELHGPTLL 600
Db 1566 THIDAHFLSOTKQAGDNFYLVAQATVCARAQAPPSWQMKCLIRLKPTELHGPTLL 1625
QY 601 YRLGAVQNEVTTHTPIKYIMACMSADLEVVTTWVLVGGVLAALAAAYCLTTGWSVIVGR 660
Db 1626 YRLGAVHNEVTLTHPVKIYIMACMSADLEVVTTWVLVGGVLAALAAAYCLTTGWSVIVGR 1685
QY 661 IILSGKPAIIPDREVLRYREDEMEECASHLPYIBQGMOLAEQFKQKAIGLLQATKQAEA 720
Db 1686 IILSGKPAIIPDREVLRYREDEMEECASHLPYIBQGMOLAEQFKQKAIGLLQATKQAEA 1745
QY 721 AAPVWESKWTLEAFWAKHWNFTSGIQLAGLSTLPCNPAIASLMAFTASITSPLTQH 780
Db 1746 AAPVWESKWTLEAFWAKHWNFTSGIQLAGLSTLPCNPAIASLMAFTASITSPLTQH 1805
QY 781 TLLFNILGGWVAQAALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 840
Db 1806 TLLFNILGGWVAQAALAPPSAASAFVAGIAGAAVGSIGLKVLDIILAGYGAGVAGALVA 1865
QY 841 FKVMSGEMPTEDLVNLLPAIIPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 900
Db 1866 FKVMSGEMPTEDLVNLLPAIIPGALVGVVCAAILRRHVGPGEVAVQWNRLLIAFASR 1925
QY 901 GNVSPHYVPESSDAARVTQILSSLTITOLLKELHOWINEDCSTPCSGMLRDVNDWIC 960
Db 1926 GNVSPHYVPESSDAARVTQILSSLTITOLLKELHOWINEDCSTPCSGMLRDVNDWIC 1985
QY 961 TVLTDEKTLQSKLLPRLPGVPPFFSCQRYGKVMRGDGMQTTCPGCAQITGHVKGSMR 1020
Db 1986 TVLTDFKTLQSKLLPRLPGVPPFFSCQRYGKVMRGDGMQTTCPGCAQISGHVKGSMR 2045
QY 1021 IVGPRICSNTHGTFPFINAYTTGCTPSPAPNYSRALMRVAABEYVEVTVRGDFHYVTGM 1080
Db 2046 IVGPKTCSNTHGTFPFINAYTTGCTPSPAPNYSRALMRVAABEYVEVTVRGDFHYVTGM 2105
QY 1081 TTDNVKCPQVPAPEPTEVDGVRHLHRYAPACKPLLRREYVELVGLNOYLVSQLPCEPE 1140
Db 2106 TTDNLKCPQVPAPEPTEVDGVRHLHRYAPACKPLLRREYVFGVLNQYLVSQLPCEPE 2165
QY 1141 PDVAVLTSMLTDFSHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTTHDSDPAD 1200
Db 2166 PDVAVLTSMLTDFSHITAEAKRLARGSPPSLASSASQLSAPSLKATCTTTHDSDPAD 2225
QY 1201 LIENALLWRQMGNTIRVSEENKVILDSFEPLOABEDREVSVPABILRRSRKFPFRAM 1260
Db 2226 LIENALLWRQMGNTIRVSEENKVILDSFELRABEDREVSVPABILRRSRKFPFRAM 2285
QY 1261 PIWAPDYNPPLLESWKDPDYVPVHVGCPLPPAKAPPPIPPRRKRTVLSSETVSSALA 1320
Db 2286 PIWAPDYNPPLLESWKDPDYVPVHVGCPLPPKGPPIPPRRKRTVLSSETVSSALA 2345
QY 1321 ELATKTFGSSSAVDSGTATAPDQPSDDGAGSDVESYSSMPPLGEPEGDPDLSGWS 1380
Db 2346 ELATKTFGSSSAVDSGTATAPDQTSDDGAGSDVESYSSMPPLGEPEGDPDLSGWS 2405
QY 1381 STVSEEBASEDVVCCSMSTWTGALITPCAAEETKLPINALNSLLRHNNLVYATTSRSAS 1440
Db 2406 STVSEEBASEDVVCCSMSTWTGALITPCAAEESKLPINALNSLLRHNNLVYATTSRSAG 2465
QY 1441 LRQKVVTFDLQVLDHYRDVLKEMKAKASTVXAKULLSVEEACKLTPPHSAKSRFGYGAK 1500
```

```
Db 2466 LRQKVVTFDLQVLDHYRDVLKEMKAKASTVXAKULLSVEEACKLTPPHSAKSRFGYGAK 2525
QY 1501 DVRLNLSKAVNHRSVWKDLLEDTETPIDTTIMAKNEVFCVQPKGGRKPARLIVFPDLG 1560
Db 2526 DVRLNLSKAVNHRSVWKDLLESETPIDTTIMAKNEIFCVQPKGGRKPARLIVFPDLG 2585
QY 1561 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPGQRVEFLVNAWKAKKCPMGFPAYDTRCFDS 1620
Db 2586 VRVCEKMALYDVVSTLPOAVMGSSYGFQYSPGQRVEFLVNTWKKCPMGFSYDTRCFDS 2645
QY 1621 TVTENDIRIBESYIQCDDLAPEARQAIRSLTERLYIGPLTNSKGQNGYRRCRASGVLT 1680
Db 2646 TVTENDIRIBESYIQCDDLAPEARQAIRSLTERLYVGGPLTNSKGQNGYRRCRASGVLT 2705
QY 1681 TSCGNTLTCVLAACAAACRAAKLOCTMLVCGDDLVVICESAGTQDEASLRAPTEAMTRY 1740
Db 2706 TSCGNTLTCVLAACAAACRAAKLOCTMLVNGDDLVVICESAGTQDEASLRAPTEAMTRY 2765
QY 1741 SAPPGDPPKPEYDLELITSCSSNVSVVAHDASGRVYVYLTDRPTTPLARAAMETARHTFPV 1800
Db 2766 SAPPGDPPKPEYDLELITSCSSNVSVVAHDASGRVYVYLTDRPTTPLARAAMETARHTFPV 2825
QY 1801 SWLGNIIIMYAPTLWARMILMTHFESILLAQOLEKALDCOYIYACYSIEPLDLPQIIERL 1860
Db 2826 SWLGNIIIMYAPTLWARMILMTHFESILLAQOLEKALDCOYIYACYSIEPLDLPQIIERL 2885
QY 1861 HGLSAFSLHSYSGEINRVASCLRKLVPLRVHRHARSVRARLLSQGGRATTCGKYL 1920
Db 2886 HGLSAFSLHSYSGEINRVASCLRKLVPLRVHRHARSVRARLLSQGGRATTCGKYL 2945
QY 1921 NWAVRTKLTPTIPAAASQDLSSWVAGYSGGDIYHSLSRARPRWFMWCLLLSVGVGIY 1980
Db 2946 NWAVRTKLTPTIPAAASQDLSSWVAGYSGGDIYHSLSRARPRWFMWCLLLSVGVGIY 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010

RESULT 12
Q9QIY8 PRELIMINARY; PRT; 3010 AA.
AC Q9QIY8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polypeptide.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD2-1;
RX MEDLINE=20013325; PubMed=10544098;
RA Nagayama K., Kuroaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity."
RL Virology 263:244-253 (1999).
DR EMBL; AF165047; AAD56182.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0251; PQ0251.
DR PIR; PQ0252; PQ0252.
DR PIR; PQ0254; PQ0254.
DR HSSP; P26663; IJXP.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
```



```
QY 1501 DVRLSKAVNHRSVMKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 1560
DB 2526 DVRLSKAVNHHSVMKDLLEDTEPTIDTTIMAKNEVFCVQPEKGRKPARLIVFPDLG 2585
QY 1561 VRCEKALYDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWAKKCPMGFAYDTRCFDS 1620
DB 2586 VRCEKALYDVVSTLPQAVMGSSYGFQYSPGQRVEFLVNAWAKKSPMGFAYDTRCFDS 2645
QY 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLNSKQNGYRCRASGVLT 1680
DB 2646 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYIGGLPLNSKQNGYRCRASGVLT 2705
QY 1681 TSCGNTLTLCYKAAACRAAKLQDCTMLVCGDDLWVICSAGTORDEASLRASFTAMTRY 1740
DB 2706 TSCGNTLTLCYKASACRAAKLQDCTMLVCGDDLWVICSAGIQEDASLRVFTAMTRY 2765
QY 1741 SAPGDPKPEYDLELITSCSSNSVVAHDASGKRVYVLTDRPTTPLARAAMETARHTPVN 1800
DB 2766 SAPGDPKPEYDLELITSCSSNSVVAHDASGKRVYVLTDRPTTPLARAAMETARHTPIN 2825
QY 1801 SWLGNIIIMYAPTLWARMLTMHTFSSILLAQLEKALDCQIYGACYSIEPLDLFQIQLR 1860
DB 2826 SWLGNIIIMYAPTLWARMLTMHTFSSILLAQLEKALDCQIYGATYSIEPLDLFQIQLR 2885
QY 1861 HGLSFAFSLHSYSPGEINRVASCLRKLGVPPLRVWHRARSVRALLSQGGAATCGKYL 1920
DB 2886 HGLSFAFSLHSYSPGEINRVASCLRKLGVPPLRVWHRARSVRALLSQGGAATCGKYL 2945
QY 1921 NNAVTKLKTPTIPAAQOLDSSWFFVAGYSGDDIYHLSLRARPRFWMCCLLLSVGVGIY 1980
DB 2946 NNAVTKLKTPTIPAAQOLDSSWFFVAGYSGDDIYHLSLRARPRFWMCCLLLSVGVGIY 3005
QY 1981 LLEPNR 1985
DB 3006 LLEPNR 3010

RESULT 13
P90193 PRELIMINARY; PRT: 3010 AA.
ID P90193 AC
AC P90193;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RN SEQUENCE FROM N.A.
RC STRAIN=HCV-1b;
RC MEDLINE=95340824; PubMed=7542279;
RA Enomoto N., Sakuma I., Asahina Y., Kurosaki M., Murakami T.,
RA Yamamoto C., Izumi N., Marumo F., Sato C.;
RT "Comparison of full-length sequences of interferon-sensitive and
RT resistant hepatitis C virus 1b.";
RN J. Clin. Invest. 96:224-230 (1995).
RN [2]
RN SEQUENCE FROM N.A.
RC STRAIN=HCV-1b;
RC Enomoto N.;
RL Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL: D50485, BAA03076.1; -.
DR PIR: A61196; A61196.
DR PIR: PQ0246; PQ0246.
DR PIR: PS0329; PS0329.
DR HSSP: P26663; IUXP.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0019028; C:viral capsid; IEA.
DR GO: GO:0019031; C:viral envelope; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0008026; F:ATP-dependent helicase activity; IEA.
```

```
DR GO: GO:0003723; F:RNA binding; IEA.
DR GO: GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO: GO:0008236; F:serine-type peptidase activity; IEA.
DR GO: GO:0005198; F:structural molecule activity; IEA.
DR GO: GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO: GO:0006350; P:transcription; IEA.
DR GO: GO:0019079; P:viral genome replication; IEA.
DR GO: GO:0019087; P:viral transformation; IEA.
DR InterPro: IPR000345; CytC_heme_BS.
DR InterPro: IPR001410; DEAD.
DR InterPro: IPR002522; HCV capsid.
DR InterPro: IPR002521; HCV core.
DR InterPro: IPR002519; HCV env.
DR InterPro: IPR002531; HCV NS1.
DR InterPro: IPR000745; HCV NS4a.
DR InterPro: IPR001490; HCV NS4b.
DR InterPro: IPR002868; HCV NS5a.
DR InterPro: IPR002166; HCV RdRP.
DR InterPro: IPR001650; Helicase_C.
DR InterPro: IPR004109; Peptidase_S29.
DR InterPro: IPR009003; Pept_Ser_Cys.
DR InterPro: IPR002518; Pept_U39_HCV_NS2.
DR InterPro: IPR007095; RNA_pol_DS_PS.
DR InterPro: IPR007094; RNA_pol_PSVir.
DR Pfam: PF01543; HCV capsid; 1.
DR Pfam: PF01542; HCV core; 1.
DR Pfam: PF01539; HCV env; 1.
DR Pfam: PF01560; HCV NS1; 1.
DR Pfam: PF01538; HCV NS2; 1.
DR Pfam: PF02907; HCV NS3; 1.
DR Pfam: PF01006; HCV NS4a; 1.
DR Pfam: PF01001; HCV NS4b; 1.
DR Pfam: PF01506; HCV NS5a; 1.
DR Pfam: PF00271; Helicase_C; 1.
DR Pfam: PF00998; Viral RdRP; 1.
DR SMART: SM00487; DEXDC; 1.
DR PROSITE: PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT CHAIN 810 1026 NS2.
FT CHAIN 1027 1657 NS3.
FT CHAIN 1658 1711 NS4a.
FT CHAIN 1 191 core protein.
FT CHAIN 1712 1972 NS4b.
FT CHAIN 1973 2419 NS5a.
FT CHAIN 2420 3010 NSSB.
FT CHAIN 192 383 E1.
FT CHAIN 384 809 E2.
FT CHAIN 3010 AA; 326886 MW; 21CD35B3DAC02B84 CRC64;
SQ SEQUENCE 3010 AA; 326886 MW; 21CD35B3DAC02B84 CRC64;

Query Match 97.3%; Score 10182; DB 2; Length 3010;
Best Local Similarity 96.6%; Pred. No. 0;
Matches 1918; Conservative 34; Mismatches 33; Indels 0; Gaps 0;

QY 1 MAPIYASQOTRGLGCIITSLTGRBNQVEGVVYVSTATQSFATCNGVCTVYHGA 60
DB 1026 LAPIYASQOTRGLGCIITSLTGRDNQVEGVVYVSTATQSFATCNGVCTVYHGA 1085
QY 61 GSKTLGPKGPIQTQMTNVNDQDLVGMQAPPGARSLTPCTCGSSDLVLTTHADVIPVRR 120
DB 1086 GSKTLGSKGPIQTQMTNVNDQDLVGMQAPPGARSLTPCTCGSSDLVLTTHADVIPVRR 1145
QY 121 GDSRGLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180
DB 1146 GDGRGLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVSMETTM 1205
QY 181 RSPVFTDNSPPAVPQTFQVAHLHAPTGSCKSTKVPAAVAAQGYKVLNPNVVAATLFGF 240
DB 1206 RSEVFTDNSPPAVPQTFQVAHLHAPTGSCKSNKVPVEYAAQGYKVLNPNVVAATLFGF 1265
QY 241 AYMSKAHGIDPNIRTCVTRITTGAPITYTYGKFLADGGCGSGAYDIICDCCHSTDST 300
DB 1266 AYMSKAHGVDPNIRTCVTRITTGAPITYTYGKFLADGGCGSGAYDIICDCCHSTDST 1325
```

```
QY 301 ILIGTIVLQDQATAGARLWVLAATATPPGSVTVPHENIEEVALSSTGEBIPFYKAIPIETI 360
Db 1326 ILIGTIVLQDQATAGARLWVLAATATPPGSVTVPHENIEEVALSNTGEIPFYKAIPIETI 1385
QY 361 KGRHLIFCHSKKKDELAALKSGILGNVAYVYRGLDVSVIPTSQDVIWVADALMTGPT 420
Db 1386 KGRHLIFCHSKKKDELAALKSGILGNVAYVYRGLDVSVIPASGDVVVWVADALMTGPT 1445
QY 421 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRQRGRTRGMRGIYRFTVPG 480
Db 1446 GDFDSVIDCNTCVTQTVDFSLDPTFTIETTTVPQDAVSRQRGRTRGMRGIYRFTVPG 1505
QY 481 ERPSGMFDSVLCBCVDACAWYELTPAETSVRLRAYLNTPLGVPCQDHLRFWESVFTGL 540
Db 1506 ERPSGMFDSVLCBCVDACAWYELTPAETSVRLRAYLNTPLGVPCQDHLRFWESVFTGL 1565
QY 541 THIDAFHLSQTKOAGDNFFYLVAQATVCARAQAPPPSDMWKCLIRLKPILHGTPLL 600
Db 1566 THIDAFHLSQTKOAGDNFFYLVAQATVCARAQAPPPSDMWKCLIRLKPILHGTPLL 1625
QY 601 YRLGAVQNEVTTHTPTTKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIVGR 660
Db 1626 YRLGAVQNEVTTHTPTTKYIMACMSADLEVVTSTWLVGGVLAALAAAYCLTTGSSVVIVGR 1685
QY 661 IILSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEOFKQKALGILQATKQAEA 720
Db 1686 IILSGKPAIIPREVLYREFDEMEECASHLPYIEQGMQLAEOFKQKALGILQATKQAEA 1745
QY 721 AAPVVEKWRITLFAFWAKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOH 780
Db 1746 AAPVVEKWALETFAWKHWNFIISGIQYLAGLSTLPGNPAIASLMAFTASITSPITTOH 1805
QY 781 TLLFNILGCVAAQALAPPAAAFVAGAGIAGAAVSGIGLKVLDVILAGYGAGVAGALVA 840
Db 1806 TLLFNILGCVAAQALAPPAAAFVAGAGIAGAAVSGIGLKVLDVILAGYGAGVAGALVA 1865
QY 841 FKVMSEMPSTEDLVNLLPAILSPGALVGVWCAAILRHHVGPGECAVOMNELLIAFASR 900
Db 1866 FKVMSEMPSTEDLVNLLPAILSPGALVGVWCAAILRHHVGPGECAVOMNELLIAFASR 1925
QY 901 GNVHSTHYVPESDAAARVTQILSSLTITQLKRLHQWINECDSTPCSGSWLRDVMWDMIC 960
Db 1926 GNVHSTHYVPESDAAARVTQILSSLTITQLKRLHQWINECDSTPCSGSWLRDVMWDMIC 1985
QY 961 TVLTDFTKWLQSKLLDRLPLGVPPFFSCQRYKGVWVRGDMQITTCPCQAQITGHVKGSMR 1020
Db 1986 TVLTDFTKWLQSKLLDRLPLGVPPFFSCQRYKGVWVRGDMQITTCPCQAQITGHVKGSMR 2045
QY 1021 IVGPRTCSTNTHGCTEPINAVTTGCTPSPAPNYSRALRWVAEYVEVTRVGVDFHVTGM 1080
Db 2046 IVGPKTCSTNTHGCTEPINAVTTGCTPSPAPNYSRALRWVAEYVEVTRVGVDFHVTGM 2105
QY 1081 TTDNVKPCQVPAPPEFFTEVDGVRHLRYAPACKPLLRREEVTFVLGNQYLVGSQLPCEPE 1140
Db 2106 TTDNLKPCQVPAPPEFFKELDGVRLHRYAPASKPLLRDEVTFOVLGNQYVVGSQLPCEPE 2165
QY 1141 PDVAVLTSMTLTDPSHTTAETAKRLRLARGSPPSLAASSASQLSAPSLKATCTTHDSDPAD 1200
Db 2166 PDVAVLTSMTLTDPSHTTAETAKRLRLARGSPPSLAASSASQLSAPSLKATCTTHDSDPAD 2225
QY 1201 LIEANILLWROEMGNITRVSEKNVTLDSFEPLOAEDEEREVSVAEILRRSRKTPPRAM 1260
Db 2226 LIEANILLWROEMGNITRVSEKNVTLDSFEPDRAEEREVSVAEILRRSRKTPPRAM 2285
QY 1261 PIWARPNDYNPLLESWKDPDYPVPPVHVCGLPFAKAPPIPPPRKRTVWLSSTVSALA 1320
Db 2286 PIWARPNDYNPLLESWKDPDYPVPPVHVCGLPFAKAPPIPPPRKRTVWLSSTVSALA 2345
QY 1321 ELATKTPGSSSASVDSGTATASDPQSDSDGAGSVESYSSPPLGEFGDPLSDGSW 1380
Db 2346 ELATKTPGSSSASVDSGTATASDPQSDSDGAGSVESYSSPPLGEFGDPLSDGSW 2405
```

```
1381 STVSEASEDDVCCSMSTWTGALITPCAABETKLPINALSNSLRHHNLVYATTSSAS 1440
Db 2406 STVSEASEDDVCCSMSTWTGALITPCAABESKLPINALSNSLRHHNMVYATTSSAS 2465
QY 1441 LRQKXVTFRDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTTPPHSARSKFGYGA 1500
Db 2466 QRQKXVTFRDRLQVLDHRYDVLKEMKAKASTVKAKLLSVEEACKLTTPPHSARSKFGYGA 2525
QY 1501 DYNRLSSKAVNHIRSVWKDLEDTEPTDITTIMAKNEVFCVQPEKGRKPARLIVPPDLG 1560
Db 2526 DYNRLSSKAVNHIRSVWKDLEDTEPTDITTIMAKNEVFCVQPEKGRKPARLIVPPDLG 2585
QY 1561 VVCEKMALYDVVSTLPOAVMGSSYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDS 1620
Db 2586 VVCEKMALYDVVSTLPOAVMGSPYGFQYSPQORVEFLVNAWKAKKCPMGFAYDTRCFDS 2645
QY 1621 TVTENDIRVEESIYOCCLDAPARQAIRSLTRLYTGGPLTNSKQONCGYRRCRASGLT 1680
Db 2646 TVTESDIRVEESIYOCCLDAPARQAIRSLTRLYTGGPLTNSKQONCGYRRCRASGLT 2705
QY 1681 TSCGNLTLCYLAACAAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTEAMTRY 1740
Db 2706 TSCGNLTLCYLAACRAAKLQDCTMLVCGDDLVVICESAGTQDEASLRAFTEAMTRY 2765
QY 1741 SAPPDGPPEYDLELITSCSSNVSAHDASGRVYVYLTTRDPTTPLARAWEATARTPVN 1800
Db 2766 SAPPDGPPEYDLELITSCSSNVSAHDASGRVYVYLTTRDPTTPLARAWEATARTPVN 2825
QY 1801 SWLGNIMVAPTILWARMILMTHFFSILLAQOELEKALDCQIYGACYSIEPLDLPQIIORL 1860
Db 2826 SWLGNIMVAPTILWARMILMTHFFSILLAQOELEKALDCQIYGACYSIEPLDLPQIIORL 2885
QY 1861 HGLSAFSLHSYSGEINRVASCLRKLGVPLRVWRHRSVRARLSSQGGRAATCGKLYF 1920
Db 2886 HGLSAFSLHSYSGEINRVASCLRKLGVPLRVWRHRSVRARLSSQGGRAATCGKLYF 2945
QY 1921 NNAVRTKLTLPAAASQDLSWVAGYSGGDIYHLSRARPRWPMWMLLLSVGVGIY 1980
Db 2946 NNAVRTKLTLPAAASQDLSWVAGYSGGDIYHLSRARPRWPMWMLLLSVGVGIY 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010
```

```
RESULT 14
Q9J3H3 PRELIMINARY; PRT; 3010 AA.
AC Q9J3H3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
CX NCBI_TaxID=11103;
RN [1]_TaxID=11103;
RP SEQUENCE FROM N.A.
RC STRAIN=MD19;
RA Nagayama K., Kurosaki M., Enomoto N., Miyasaka Y., Marumo F., Sato C.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
EMBL: AF207760; AAF65950.1; -.
PIR: A61196; A61196.
FIR: PS0329; PS0329.
DR HSP; P26663; IUXP.
DR GO:0016021; C:integral to membrane; IEA.
DR GO:0019028; C:viral capsid; IEA.
DR GO:0019031; C:viral envelope; IEA.
DR GO:0005524; F:ATP binding; IEA.
DR GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO:0003723; F:RNA binding; IEA.
DR GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
```



DR GO: 0008236; F: serine-type peptidase activity; IEA.  
DR GO: 0005198; F: structural molecule activity; IEA.  
DR GO: 0006508; P: proteolysis and peptidolysis; IEA.  
DR GO: 0006350; P: transcription; IEA.  
DR GO: 0019079; P: viral genome replication; IEA.  
DR GO: 0019087; P: viral transmembrane; IEA.  
DR InterPro: IPR000345; CytC heme\_BS.  
DR InterPro: IPR001410; DEAD.  
DR InterPro: IPR002522; HCV capsid.  
DR InterPro: IPR002521; HCV core.  
DR InterPro: IPR002519; HCV env.  
DR InterPro: IPR002531; HCV NS1.  
DR InterPro: IPR000745; HCV NS4a.  
DR InterPro: IPR001490; HCV NS4b.  
DR InterPro: IPR002868; HCV NS5a.  
DR InterPro: IPR002166; HCV RdRP.  
DR InterPro: IPR001650; Helicase C.  
DR InterPro: IPR004109; Peptidase S29.  
DR InterPro: IPR009003; Pept Ser Cys.  
DR InterPro: IPR002518; Pept\_U33 HCV NS2.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR Pfam: PF01543; HCV capsid; 1.  
DR Pfam: PF01542; HCV core; 1.  
DR Pfam: PF01539; HCV env; 1.  
DR Pfam: PF01560; HCV NS1; 1.  
DR Pfam: PF01538; HCV NS3; 1.  
DR Pfam: PF02907; HCV NS3; 1.  
DR Pfam: PF01006; HCV NS4a; 1.  
DR Pfam: PF01001; HCV NS4b; 1.  
DR Pfam: PF01506; HCV NS5a; 1.  
DR Pfam: PF00271; Helicase C; 1.  
DR Pfam: PF00998; Viral RdRP; 1.  
DR SMART; SMO0487; DEXDC; 1.  
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.  
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;  
KW Polyprotein; Transmembrane.  
SQ SEQUENCE 3010 AA; 327235 MW; 44C34677649CB8DD CRC64;  
  
Query Match 97.2%; Score 10177; DB 2; Length 3010;  
Best Local Similarity 96.1%; Pred. No. 0;  
Matches 1908; Conservative 44; Mismatches 33; Indels 0; Gaps 0;  
  
QY 1 MAPITAYSOQTRGLLCITSLTGRDNQVEGVQVSTATQSFATCNGVCWTVYHGA 60  
DB 1026 LAPITAYSOQTRGLLCITSLTGRDNQVEGVQVSTATQSFATCNGVCWTVYHGA 1085  
  
QY 61 GSXKLAGPKGPIQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLVLTVRHADVIPVRR 120  
DB 1086 GAKTLAGPKGPIQMYTNVDQDLVGWQAPPGARSLTPCTCGSSDLVLTVRHADVIPVRR 1145  
  
QY 121 GDSRGLLSPRPVSYLKGSGGPGLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180  
DB 1146 GDSRGLLSPRPVSYLKGSGGPGLLCPGSHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 1205  
  
QY 181 RSPVFTDSSPPAVPOTFOVAHLHAPTGSCKSTKVAAYAAQYKVLVLPNSVAATLFG 240  
DB 1206 RSPVFTDSSPPAVPOTFOVAHLHAPTGSCKSTKVAAYAAQYKVLVLPNSVAATLFG 1265  
  
QY 241 AYMSKAHGIDPNIRTVRTTGTAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 300  
DB 1266 AYMSKAHGIEPRTVTRITGTAPITYSTYKFLADGGCGGAYDIIICDECHSDSTT 1325  
  
QY 301 ILGIGTVLQOATAGARLVVLTATPPGVSVPVPHNIEEVALSSCTGEIPFYKAIPIETI 360  
DB 1326 ILGIGTVLQOATAGARLVVLTATPPGVSVPVPHNIEEVALSSCTGEIPFYKAIPIETI 1385  
  
QY 361 KGRGHLIFCHSKKKCKDELAAGSLGNAVAYVYRGLDVSVIPISGVIVVATDALTMTGT 420  
DB 1386 KGRGHLIFCHSKKKCKDELAAGSLGNAVAYVYRGLDVSVIPISGVIVVATDALTMTGT 1445  
  
QY 421 GQFDSVIDCNTCTVQTVDFSLDPTFTTETTTVPQDAVSRQRGRTRGRMGRIYRFTPG 480  
DB 481 ERPSGMFDDSSVLCBCEYDAGCAWYELTAPETSRLRAYLNTPLPVCODHLEFVESVFTGL 540  
DB 1506 ERPSGMFDDSSVLCBCEYDAGCAWYELTAPETSRLRAYLNTPLPVCODHLEFVESVFTGL 1565  
QY 541 THIDAHFLSQTQKAGDNFPYLVAQATVCAQAAPPSPSWDMKCLRLKPTLHGPTPL 600  
DB 1566 THIDAHFLSQTQKAGDNFPYLVAQATVCAQAAPPSPSWDMKCLRLKPTLHGPTPL 1625  
QY 601 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVVTWLVGVGLAALAAAYCLTGTGWSVLVGR 660  
DB 1626 YRLGAVQNEVTTTHPTTKYIMACMSADLEVVVTWLVGVGLAALAAAYCLTGTGWSVLVGR 1685  
QY 661 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLAEQFKQKALGLQTATQAEA 720  
DB 1686 IILSGKPAIIPDREVLYREFDEMEECASHLPYIEQGMQLAEQFKQKALGLQTATQAEA 1745  
QY 721 AAPVSESKRTLEAFWAKHWNFI SGIQYLAGLSTLPGNPAIASLMFTAFTSPLTQH 780  
DB 1746 AAPVSESKRTLEAFWAKHWNFI SGIQYLAGLSTLPGNPAIASLMFTAFTSPLTQH 1805  
QY 781 TLLFNILGGWVAQAAPPASAASAFVAGTAGAAGVSGTGLGKVLVDIILAGYAGVAGALVA 840  
DB 1806 TLLFNILGGWVAQAAPPASAASAFVAGTAGAAGVSGTGLGKVLVDIILAGYAGVAGALVA 1865  
QY 841 FKVMGEMFSTEDLVNLLPAIFAILSPGALVGVVCAAILRRHVGPGEVAVMMNLLIAPASR 900  
DB 1866 FKVMGEMFSTEDLVNLLPAIFAILSPGALVGVVCAAILRRHVGPGEVAVMMNLLIAPASR 1925  
QY 901 GNVHSTHVVPESDAARVQIILSSLTITQLLKHQWNEDECSTPCSGSWLDRVMDWIC 960  
DB 1926 GNVHSTHVVPESDAARVQIILSSLTITQLLKHQWNEDECSTPCSGSWLDRVMDWIC 1985  
QY 961 TVLTDFTKLQSKLLPRLGVPPFSCORGKGVWGDGMOTTCGCAOITGHVKNKSMR 1020  
DB 1986 TVLADFTKLQSKLLPRLGVPPFSCORGKGVWGDGMOTTCGCAOITGHVKNKSMR 2045  
QY 1021 IVPRTCSNTWHGTFFINAYTTGPTCPSPAPNYSRALWRVAABEYVEVTVGVPHYVTGM 1080  
DB 2046 IVPRTCSNTWHGTFFINAYTTGPTCPSPAPNYSRALWRVAABEYVEVTVGVPHYVTGM 2105  
QY 1081 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLLEEVTVFLVGNQYLVGSQLPCEPE 1140  
DB 2106 TTDNVKCPQVPAPEFTEVDGVLHRYAPACKPLLEEVTVFLVGNQYLVGSQLPCEPE 2165  
QY 1141 PDVAVLTSLMTDPSHITTAETAKRLARGSPPSLIASSASQLSAPSKATCTTHDSDPAD 1200  
DB 2166 PDVAVLTSLMTDPSHITTAETAKRLARGSPPSLIASSASQLSAPSKATCTTHDSDPAD 2225  
QY 1201 LIEANLLWQEMGNNITRVESKNVILDSFEPLQAEDEEVSVPAEILRKRKFPFRAM 1260  
DB 2226 LIEANLLWQEMGNNITRVESKNVILDSFEPLQAEDEEVSVPAEILRKRKFPFRAM 2285  
QY 1261 PIWARPDPNPPLESKWDYPPVWHGCPPLPAKAPPIPPPRKRRTVLSSTVSSALA 1320  
DB 2286 PIWARPDPNPPLESKWDYPPVWHGCPPLPAKAPPIPPPRKRRTVLSSTVSSALA 2345  
QY 1321 ELATKTFGSSSADVSDGTATASPOQSDGAGSDVESYSSMPPLEGGPDLSDGWS 1380  
DB 2346 ELATKTFGSSSADVSDGTATASPOQSDGAGSDVESYSSMPPLEGGPDLSDGWS 2405  
QY 1381 STVSEASDDVCCSMSTYTWGALITPCAAEETKLPINALNSLLRHHNVVATTSRSAS 1440  
DB 2406 STVSEASDDVCCSMSTYTWGALITPCAAEETKLPINALNSLLRHHNVVATTSRSAS 2465  
QY 1441 LRQKVTDFRLQVLDHHDVLEKEMAKASTVAKLLSVEEACKLTPPHSARSKFYGAK 1500  
DB 2466 LRQKVTDFRLQVLDHHDVLEKEMAKASTVAKLLSVEEACKLTPPHSARSKFYGAK 2525  
QY 1501 DVNRLSSKAVNHRSWKDLLEDTEPTIDTTIMAKNEVFCVQPKGRKPARLIVFPDLG 1560  
DB 2526 DVNRLSSKAVNHRSWKDLLEDTEPTIDTTIMAKNEVFCVQPKGRKPARLIVFPDLG 2585

```
QY 1561 VRVCEKALYDVVSTLPQAVNGSSYGFQYSPQQRVEFLVNAWAKKCPMGFAYDTRCFDS 1620
Db 2586 VRVCEKALYDVVSTLPQAVNGASYGFQYSPQQRVEFLVNAWKKCPMGFSYDTRCFDS 2645
QY 1621 TVTENDIIVERISYOCCLAPEARQAIISLTERLYIGGLPLTNSKQNGCYRRCRASGVLT 1680
Db 2646 TVTENDIITERISYOCCLAPARQAIISLTERLYIGGLPLTNSKQNGCYRRCRASGVLT 2705
QY 1681 TSCGNLTLCYLKAAACRAAKLQDCTMLVCGDDLVIICESAGTQDEASLRAFTAMTRY 1740
Db 2706 TSCGNLTLCYLKATAACRAAKLQDCTMLVCGDDLVIICESAGTQDEAANLRFTEAMTRY 2765
QY 1741 SAPPGDPKPEYDLELITSCSSNVAVHDAAGSKRYVYLTRDPTTFLARAAMETARTHPVN 1800
Db 2766 SAPPGDPKPEYDLELITSCSSNVAVHDAAGSKRYVYLTRDPTTFLARAAMETARTHPVN 2825
QY 1801 SWLGNIMYAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPQIIQRL 1860
Db 2826 SWLGNIMYAPTLWARMILMTHFFSILLAQEQLKALDCQIYGACYSIEPLDLPQIIQRL 2885
QY 1861 HGLSAFSLHSYSPGBINRVASCLRLKGLVPPLEVRHRSVRARLLSQGRAATCGKYL 1920
Db 2886 HGLSAFSLHSYSPGBINRVASCLRLKGLVPPLEVRHRSVRARLLSQGRAATCGKYL 2945
QY 1921 NWAETKLTLPIDPAASQDLSSWFVAGYSGGDIYHLSRARPFRMFWMLLLLSVGVGII 1980
Db 2946 NWAETKLTLPIDPAAYQLDLSGWFVAGYSGGDIYHLSVRARPFRMFWMLLLLSVGVGII 3005
QY 1981 LLPNR 1985
Db 3006 LLPNR 3010

RESULT 15
Q9QIX6 PRELIMINARY; PRT; 3010 AA.
AC Q9QIX6;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Polyprotein.
OS Hepatitis C virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
OC Hepacivirus.
OX NCBI_TaxID=11103;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MD8-1;
RX MEDLINE=20013325; PubMed=10544098;
RA Nagayama K., Kurosaki M., Enomoto N., Maekawa S.y., Miyasaka Y.,
RA Tazawa J.i., Izumi N., Marumo F., Sato C.;
RT "Time-related changes in full-length hepatitis C virus and hepatitis
RT activity.";
RL Virology 263:244-253(1999).
DR EMBL; AF165059; AAD56194.1; -.
DR PIR; A61196; A61196.
DR PIR; PQ0246; PQ0246.
DR PIR; PQ0254; PQ0254.
DR PIR; PQ0804; PQ0804.
DR PIR; PS0329; PS0329.
DR HSP; P26663; 1JXP.
DR MEROPS; S29.002; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0019028; C:viral capsid; IEA.
DR GO; GO:0019031; C:viral envelope; IEA.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0008026; F:ATP-dependent helicase activity; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003968; F:RNA-directed RNA polymerase activity; IEA.
DR GO; GO:0008236; F:serine-type peptidase activity; IEA.
DR GO; GO:0005198; F:structural molecule activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
GO; GO:0006350; P:transcription; IEA.
GO; GO:0019079; P:viral genome replication; IEA.
GO; GO:0019087; P:viral transformation; IEA.
DR InterPro; IPR000345; CytC_heme_BS.
DR InterPro; IPR001410; DEAD.
DR InterPro; IPR002522; HCV capsid.
DR InterPro; IPR002521; HCV core.
DR InterPro; IPR002519; HCV env.
DR InterPro; IPR002531; HCV NS1.
DR InterPro; IPR000745; HCV NS4a.
DR InterPro; IPR001490; HCV NS4b.
DR InterPro; IPR002868; HCV NS5a.
DR InterPro; IPR002166; HCV RdRp.
DR InterPro; IPR001650; Helicase_C.
DR InterPro; IPR004109; Peptidase_S29.
DR InterPro; IPR009003; Pept_Ser_Cys.
DR InterPro; IPR002518; Pept_U39_HCV_NS2.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR007094; RNA_pol_Psivir.
DR Pfam; PF01543; HCV capsid; 1.
DR Pfam; PF01542; HCV core; 1.
DR Pfam; PF01539; HCV env; 1.
DR Pfam; PF01560; HCV NS1; 1.
DR Pfam; PF01538; HCV NS2; 1.
DR Pfam; PF02907; HCV NS3; 1.
DR Pfam; PF01006; HCV NS4a; 1.
DR Pfam; PF01001; HCV NS4b; 1.
DR Pfam; PF01506; HCV NS5a; 1.
DR Pfam; PF00271; Helicase_C; 1.
DR Pfam; PF00998; Viral_RdRp; 1.
DR SMART; SM00487; DEXDC; 1.
DR PROSITE; PS00190; CYTOCHROME C; UNKNOWN 1.
KW Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
SQ SEQUENCE 3010 AA; 327468 MW; 4613744EC4D4A013 CRC64;
Query Match 97.2%; Score 10177; DB 2; Length 3010;
Best Local Similarity 96.5%; Pred. No. 0;
Matches 1916; Conservative 41; Mismatches 28; Indels 0; Gaps 0;
QY 1 MAPITAYSQOTRGLGCIITSLTGRDRNQVEGEVQVWSTATQSFATCNGVCWTYYHGA 60
Db 1026 LAPITAYSQOTRGLGCIITSLTGRDRKNQVEGEVQVWSTATQSFATCNGVCWTYYHGA 1085
QY 61 GSKTLAGPKGITOMYTNVDQDLVQWQAPPGARSLTPTCTGSSDLVLTTHADVIPVRR 120
Db 1086 GSKTLAGPKGITOMYTNVDQDLVQWQAPPGARSLTPTCTGSSDLVLTTHADVIPVRR 1145
QY 121 GDSRGLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFVPVSMETTM 180
Db 1146 GDGRGLSPRPVSYLKGSSGGPLLCPSGHAVGIFRAAVCTRGVAKAVDFIPVSMETTM 1205
QY 181 RSPVFTDNSSPPAVPQTQVAHLHAPTSGSKSTKVPAAVAAQGYKVLVLPNSVAATLGG 240
Db 1206 RSPVFTDNSSPPAVPQTQVAHLHAPTSGSKSTKVPAAVAAQGYKVLVLPNSVAATLGG 1265
QY 241 AYMSKAHGIDPNIRTVGRTITTTGAPITYSTYKGLADGCGSGAGVDIIICDCHSDTST 300
Db 1266 AYMSKAYGTDPNIRTVGRTITTTGAPITYSTYKGLADGCGSGAGVDIIICDCHSDTST 1325
QY 301 ILGIGTVLDQAEAGARLVVLTATPPGSSVTVPHENIEVALSSTGEIPFFYKALPIETI 360
Db 1326 ILGIGTVLDQAEAGARLVVLTATPPGSSVTVPHENIEVALSSTGEIPFFYKALPIETI 1385
QY 361 KGRHLIFCHSKKKDELAALKSLGLNAVAYRGLDVSIVPTSGDVIVVATDALMTGT 420
Db 1386 KGRHLIFCHSKKKDELAALKSLGLNAVAYRGLDVSIVPTSGDVIVVATDALMTGT 1445
QY 421 GPFDSVIDCNTCVTQTVDFSLDPTFTTETTTVPQAVSRQRGRGTMGYRFTVTPG 480
Db 1446 GPFDSVIDCNTCVTQTVDFSLDPTFTTETTTVPQAVSRQRGRGTMGYRFTVTPG 1505
QY 481 ERPSGMFDSSVLCECYDAGCAWYELTPAETSRLRAYLNTPLGLPVCQDHLERFESVFTGL 540
```

Db 1506 ERPSGNFSDSVLCECYDAGCANYELTPTAETSVRLRAYLNTPLGVPUCQDHLHFWESVFTGL 1565  
QY 541 THIDAHFLSQTQAGDNFPYVAYQATVCARAQPPSWDMWKCLIRLKPTLHGPTPL 600  
Db 1566 THIDAHFLSQTQAGDNFPYVAYQATVCARAQPPSWDMWKCLIRLKPTLHGPTPL 1625  
QY 601 YRLGATQNEVTHPTITKYMCMASDLVVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 660  
Db 1626 YRLGATQNEVTHPTITKYMCMASDLVVTSTWLVGGVLAALAAAYCLTTGSSVIVGR 1685  
QY 661 IILSGKPAIIPREVLIRFDEMEECASHLPYIEQGMQLABQFKOKAIGLLOPATKQAEA 720  
Db 1686 IILSGKPAIIPREVLIRFDEMEECASHLPYIEQGMQLABQFKOKAIGLLOPATKQAEA 1745  
QY 721 AAPVSEKVRTUEAFNAKHMWNFISGQYLAGLSTLPGNPAIASLMAFTASITPTTQH 780  
Db 1746 AAPVSEKVRTUEAFNAKHMWNFISGQYLAGLSTLPGNPAIASLMAFTASITPTTQH 1805  
QY 781 TLLFNILGGVAAQLAPPAAAFVAGTAGAAGVSGIGKVLVDILAGYGAGVAGALVA 840  
Db 1806 TLLFNILGGVAAQLAPPAAAFVAGTAGAAGVSGIGKVLVDILAGYGAGVAGALVA 1865  
QY 841 FKVMSEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVOMNRLIAFASR 900  
Db 1866 FKVMSEMPSTEDLVNLLPAILSPGALVGVVCAAILRRHVGPGEAGVOMNRLIAFASR 1925  
QY 901 GNVHSTHYVPESDAAARVTOILSSITITQLKRLHQMINECDSTPCSGSWLRDWDWIC 960  
Db 1926 GNVHSTHYVPESDAAARVTOILSSITITQLKRLHQMINECDSTPCSGSWLRDWDWIC 1985  
QY 961 TVLTDFKTLQSKLLPRLFGVFPFSCQYKGVWRGDMQITPCGQAIITHGVKNGSR 1020  
Db 1986 TVLTDFKTLQSKIMPRLPFGVFPFSCQYKGVWRGDMQITPCGQAIITHGVKNGSR 2045  
QY 1021 IVGPRTCSNTWGTGTFPINAYTTGCTPPAPNYSRALWVAABEYVEVTRVGDGFHYVTGM 1080  
Db 2046 IVGPRTCSNTWGTGTFPINAYTTGCTPPAPNYSRALWVAABEYVEVTRVGDGFHYVTGM 2105  
QY 1081 TTDNVKPCQVPAPEFTFVDGVRHLRYAPACKPLREBVTFLVGLNQYLVGSQLPCEPE 1140  
Db 2106 TTDNVKPCQVPAPEFTFVDGVRHLRYAPACKPLREBVTFLVGLNQYLVGSQLPCEPE 2165  
QY 1141 PDVAVLTSMLTDPHSHTAETAKRRLARGSPPSLASSASQLSAPSUKATCTTHDSPDAD 1200  
Db 2166 PDVAVLTSMLTDPHSHTAETAKRRLARGSPPSLASSASQLSAPSUKATCTTHDSPDAD 2225  
QY 1201 LIEANLLWRQEMGNITRVESENKVVILDSFPLQAEDEDEREVSVAEILRRSRKFPFRAM 1260  
Db 2226 LIEANLLWRQEMGNITRVESENKVVILDSFPLQAEDEDEREVSVAEILRRSRKFPFRAM 2285  
QY 1261 PIWARPDYNNPILLESWKDDYVPPVHVHGCPLPPAKAPPIPPRRKRKTVVLSSESTVSSALA 1320  
Db 2286 PIWARPDYNNPILLESWKDDYVPPVHVHGCPLPPAKAPPIPPRRKRKTVVLSSESTVSSALA 2345  
QY 1321 ELATKTFGSSSADVSDGTATASPDQSDGAGSVESYSMPPLEGEPPDLSDGSM 1380  
Db 2346 ELATKTFGSSSADVSDGTATASPDQSDGAGSVESYSMPPLEGEPPDLSDGSM 2405  
QY 1381 STVSEASDDVVCSSMSYTWGTGALITPCAABETKLPINALSNSLRHNLVYATTSSRSAS 1440  
Db 2406 STVSEASDDVVCSSMSYTWGTGALITPCAABETKLPINALSNSLRHNLVYATTSSRSAS 2465  
QY 1441 LRQKKVTFORLQVLDHXYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 1500  
Db 2466 LRQKKVTFORLQVLDHXYRDVLKEMKAKASTVKAKLLSVEEACKLTPPHSARSKFGYGAK 2525  
QY 1501 DYNRLSSKAVNHIRSVMKDLLEDTEPTIDTTIMAKNEVFCVOPEKGRKPARIIVFPDLG 1560  
Db 2526 DYNRLSSKAVNHIRSVMKDLLEDTEPTIDTTIMAKNEVFCVOPEKGRKPARIIVFPDLG 2585  
QY 1561 VRVCEKNALYDVVSTLPOAVMGSSYGFQYSPGORVEFLYNAMKAKCPMGFAYDTRCFDS 1620

Search completed: December 8, 2004, 12:32:06  
Job time : 281 secs

Db 2586 VRVCEKNALYDVVSTLPOAVMGSSYGFQYSPGORVEFLYNAMKAKCPMGFSYDTRCFDS 2645  
QY 1621 TVTENDIRVEESIYQCCDLAPEARQAIRSLTERLYTIGGPLTNSKGQNGYRRCRASGVL 1680  
Db 2646 TVTESDIRTEESIYQCCDLAPEARQAIRSLTERLYTIGGPLTNSKGQNGYRRCRASGVL 2705  
QY 1681 TSCGNLTLCYLKAAAACRAAKLODCTMLVCGDDLVI CESAGTOEDEASIRAFTEAMTRY 1740  
Db 2706 TSCGNLTLCYLKASACRAAKLODCTMLVNGDDLVI CESAGTOEDAASIRAFTEAMTRY 2765  
QY 1741 SAPPDGPPEYDLELITSCSNVSVAHDAAGKRVYLLTRDPTTPLARAAMETARHTPVN 1800  
Db 2766 SAPPDGPPEYDLELITSCSNVSVAHDAAGKRVYLLTRDPTTPLARAAMETARHTPVN 2825  
QY 1801 SWLGNIIIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIERL 1860  
Db 2826 SWLGNIIIMYAPTILWARMILMTHFFSILLAQEOLKALDCQIYGACYSIEPLDLPQIIERL 2885  
QY 1861 HGLSAPSLHSYSPGEINRVASCLKLGVPPLVWRHRARSVRARLLSQCGRATCGKYL 1920  
Db 2886 HGLSAPSLHSYSPGEINRVASCLKLGVPPLVWRHRARSVRARLLSQCGRATCGKYL 2945  
QY 1921 NNAVRTKLKLTIPPAASQLDLSGFWAGYSGGDIYHSLSRARPRWPMWCLLILSVGVGIY 1980  
Db 2946 NNAVRTKLKLTIPPAASRLDLSGFWAGYSGGDIYHSLSRARPRWPMWCLLILSVGVGIY 3005  
QY 1981 LLPNR 1985  
Db 3006 LLPNR 3010

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 20:18:55 ; Search time 32059 Seconds  
(without alignments)  
11784.441 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989

Sequence: 1 gccagcccccagattgggggc.....ggcctctctgcagatcaagt 7989

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.\*

1: gb.ba.\*

2: gb.htg.\*

3: gb.in.\*

4: gb.om.\*

5: gb.ov.\*

6: gb.pat.\*

7: gb.ph.\*

8: gb.pl.\*

9: gb.pr.\*

10: gb.ro.\*

11: gb.sts.\*

12: gb.sy.\*

13: gb.un.\*

14: gb.vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	7987.4	100.0	7989	6	AR406043
2	7987.4	100.0	7989	6	AX036255 Sequence
3	7987.4	100.0	7989	12	SSE242652
4	7987.4	100.0	10690	6	AJ242652 Hepatitis
5	7965.4	99.7	8001	6	AX739972 Sequence
6	7965.4	99.7	8001	6	AR406042
7	7965.4	99.7	8001	6	AX036254
8	7947.8	99.5	8001	12	SSE242654
9	7947.8	99.5	8001	6	AR406047
10	7946.2	99.5	8001	6	AX036259
11	7946.2	99.5	8001	6	AR406045
12	7939.8	99.4	8001	6	AX036257
13	7939.8	99.4	8001	6	AR406049
14	7610.4	95.3	12305	6	AX036261
15	7607.2	95.2	12315	6	AX937622
16	7329.4	91.7	8637	6	AX937621
17	7329.4	91.7	8637	6	AR406041
18	7329.4	91.7	8637	6	AX036253
19	7307.4	91.5	8649	6	AJ242651 Hepatitis
					AR406044

20	7307.4	91.5	8649	6	AX036256	AX036256 Sequence
21	7307.4	91.5	8649	12	SSE242653	AJ242653 Hepatitis
22	7272.2	91.0	8639	6	AR488100	AR488100 Sequence
23	7272.2	91.0	8639	6	AX472291	AX472291 Sequence
24	7271.2	91.0	8638	6	AR488121	AR488121 Sequence
25	7271.2	91.0	8638	6	AX472314	AX472314 Sequence
26	7248.2	90.7	8638	6	AR488105	AR488105 Sequence
27	7248.2	90.7	8638	6	AX472297	AX472297 Sequence
28	7247.2	90.7	8638	6	AR488122	AR488122 Sequence
29	7247.2	90.7	8638	6	AX472315	AX472315 Sequence
30	7245.2	90.7	8642	6	AR488101	AR488101 Sequence
31	7245.2	90.7	8642	6	AX472292	AX472292 Sequence
32	7244	90.7	8638	6	AR488104	AR488104 Sequence
33	7244	90.7	8638	6	AX472296	AX472296 Sequence
34	7228.2	90.5	8648	6	AR488103	AR488103 Sequence
35	7228.2	90.5	8648	6	AX472295	AX472295 Sequence
36	7221	90.4	8643	6	AR488102	AR488102 Sequence
37	7221	90.4	8643	6	AX472294	AX472294 Sequence
38	6679.8	83.6	9658	14	AB119282	AB119282 Hepatitis
39	6185.2	77.4	9605	6	AX739971	AX739971 Sequence
40	6185.2	77.4	9605	14	HCJ238799	AJ238799 Hepatitis
41	6185.2	77.4	11076	6	AR406040	AR406040 Sequence
42	6185.2	77.4	11076	6	AX036252	AX036252 Sequence
43	6167.6	77.2	11076	6	AR406048	AR406048 Sequence
44	6167.6	77.2	11076	6	AX036260	AX036260 Sequence
45	6166	77.2	11076	6	AR406046	AR406046 Sequence

#### ALIGNMENTS

RESULT 1  
AR406043  
LOCUS AR406043 7989 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 10 from patent US 6630343.  
ACCESSION AR406043  
VERSION AR406043.1 GI:40155170  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 7989)  
AUTHORS Bartenschlager,R.  
TITLE Hepatitis C virus culture system  
JOURNAL Patent: US 6630343-A 10 07-OCT-2003;  
FEATURES Location/Qualifiers  
source 1..7989  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Query Match 100.0%; Score 7987.4; DB 6; Length 7989;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7989; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
Db 1 GCCAGCCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
QY 61 TCTTCACGCAAGACGCTCTAGCCATGGCGTTAGTATAGTGTGTCGACGCTCCAGGAC 120  
Db 61 TCTTCACGCAAGACGCTCTAGCCATGGCGTTAGTATAGTGTGTCGACGCTCCAGGAC 120  
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACGGGTGAGTACACCGGAATTCGCG 180  
Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACGGGTGAGTACACCGGAATTCGCG 180  
QY 181 GACGACCGGTCTCTTTTGGATCAACCCGCTCAATGCTCGGAGATTTGGGCGTGCCCC 240  
Db 181 GACGACCGGTCTCTTTTGGATCAACCCGCTCAATGCTCGGAGATTTGGGCGTGCCCC 240  
QY 241 GCGAGACTGCTAGCCAGTAGTGTGGTGGCGAAAGGCGCTTGTGTACTGCTGTATAGG 300





QY	2461	GCCAAAGGATATAGGTGCTTCTGCTGAACCGTCCGTCGCCGCCACCCCTAGTGTTCGGG	2520	DB	3541	CAAAATGTGGAAGTGTCTCATACGGCTAAAGGCTACGCTGCACGGGCCCAACGCCCTGTGTG	3600
DB	2461	GCCCAAGGATATAGGTGCTTGTCTGAAACCGTCCGTCGCCGCCACCCCTAGTGTTCGGG	2520	QY	3601	TATAGGCTGGGAGCGGTTCAAAAACGAGGTTACTACACACACACCCCATACCAAAATACATC	3660
QY	2521	GGGTATATGTCTAAGGCAATAGTATCGACCCCTAAATCATCAGAACCGGGGTAAAGACCATC	2580	DB	3601	TATAGGCTGGGAGCGGTTCAAAAACGAGGTTACTACACACACCCCATACCAAAATACATC	3660
DB	2521	GGGTATATGTCTAAGGCAATAGTATCGACCCCTAAATCATCAGAACCGGGGTAAAGACCATC	2580	QY	3661	ATGGATCATGTCTGGCTGACCTGGAGGTCTGCAGGACCACTGGGTGCTGGTAGCGGA	3720
QY	2581	ACCAGGGTGGCCCATCACCTACTCCACCTATGCAAGTTCCTTTCGCCGACGGTGGTTC	2640	DB	3661	ATGGCATCATGTCTGGCTGACCTGGAGGTCTGCAGGACCACTGGGTGCTGGTAGCGGA	3720
DB	2581	ACCAGGGTGGCCCATCACCTACTCCACCTATGCAAGTTCCTTTCGCCGACGGTGGTTC	2640	QY	3721	GTCTTAGCAGCTCTGGCCGCGTATTTGCTGACACAGCAGCGGTGCTCATTTGGGCGAG	3780
QY	2641	TCTGGGGCGCTATGACATCATATATGTGATGAGTGCCTCACTCAACTGACTCGACCACT	2700	DB	3721	GTCTTAGCAGCTCTGGCCGCGTATTTGCTGACACAGCAGCGGTGCTCATTTGGGCGAG	3780
DB	2641	TCTGGGGCGCTATGACATCATATATGTGATGAGTGCCTCACTCAACTGACTCGACCACT	2700	QY	3781	ATCATCTTGTCCGAAAGCGCGCCATCATTCGCCACAGGGAAGTCTTTTACCGGAGTTC	3840
QY	2701	ATCTGGGCAATCGGCACAGTCTCTGACCAAGCGGAGCGGTGGAGCGCGACTCGTCGTG	2760	DB	3781	ATCATCTTGTCCGAAAGCGCGCCATCATTCGCCACAGGGAAGTCTTTTACCGGAGTTC	3840
DB	2701	ATCTGGGCAATCGGCACAGTCTCTGACCAAGCGGAGCGGTGGAGCGCGACTCGTCGTG	2760	QY	3841	GATGAGATGGAAGAGTGGCTCTCACCTCTCTTACATCGAACAGGGAATCAGCTCGCC	3900
QY	2761	CTCGCCACCGTACGCTCCGGATCGGTCAACGTCACGTCACATCMAAATCGAGAGGTG	2820	DB	3841	GATGAGATGGAAGAGTGGCTCTCACCTCTCTTACATCGAACAGGGAATCAGCTCGCC	3900
DB	2761	CTCGCCACCGTACGCTCCGGATCGGTCAACGTCACGTCACATCMAAATCGAGAGGTG	2820	QY	3901	GAAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCACCACCAAGCAGCGAGGCT	3960
QY	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCCATCGAGACCATC	2880	DB	3901	GAAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCACCACCAAGCAGCGAGGCT	3960
DB	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCCATCGAGACCATC	2880	QY	3961	GCTCTCTCCGTTGGTGAATCCAAGTGGCGGACCTCGAAGGCTTCTGGCGCAAGCATATG	4020
QY	2881	AAGGGGGGAGGCACTCATTTTCTGGCAATCCAAAGAAATGTGATGAGCTCGCGCG	2940	DB	3961	GCTCTCTCCGTTGGTGAATCCAAGTGGCGGACCTCGAAGGCTTCTGGCGCAAGCATATG	4020
DB	2881	AAGGGGGGAGGCACTCATTTTCTGGCAATCCAAAGAAATGTGATGAGCTCGCGCG	2940	QY	4021	TGGAATTTTATCAGCGGATACAAATTTATGACAGGCTTGTCCACTCTGCTTCCCAACCCC	4080
QY	2941	AAGCTGTCCGCTCCGACCTCAATGCTGACATATTAACCGGGCTTGTATCGTC	3000	DB	4021	TGGAATTTTATCAGCGGATACAAATTTATGACAGGCTTGTCCACTCTGCTTCCCAACCCC	4080
DB	2941	AAGCTGTCCGCTCCGACCTCAATGCTGACATATTAACCGGGCTTGTATCGTC	3000	QY	4081	GCGATGATCATCTGATGGGATTCACAGCTTATACAGGCTTGTCCACTCTGCTTCCCAACCCC	4140
QY	3001	ATACCAACTAGCGGAGAGTCAATGTCTGACCAACGAGCGCTTAATGACGGGCTTTACC	3060	DB	4081	GCGATGATCATCTGATGGGATTCACAGCTTATACAGGCTTGTCCACTCTGCTTCCCAACCCC	4140
DB	3001	ATACCAACTAGCGGAGAGTCAATGTCTGACCAACGAGCGCTTAATGACGGGCTTTACC	3060	QY	4141	ACCTCTCTGTTTAACTCTCTGGGGGATGGTGGCGCCCAACTTGTCTTCCCAAGCGCT	4200
QY	3061	GGCAATTTGCACTCAGTGATGATGCAATATGATGTGACCCAGACAGTCACTTACG	3120	DB	4141	ACCTCTCTGTTTAACTCTCTGGGGGATGGTGGCGCCCAACTTGTCTTCCCAAGCGCT	4200
DB	3061	GGCAATTTGCACTCAGTGATGATGCAATATGATGTGACCCAGACAGTCACTTACG	3120	QY	4201	GCTTCTCTCTTCTGATGAGCGCGGATCGCTGGAGCGGCTGTGGCAGCATAGGCTTGGG	4260
QY	3121	CTGACCCGACCTTACCATGACAGCAGCCGTCGCAACAGCGGCTGTGACGTCG	3180	DB	4201	GCTTCTCTCTTCTGATGAGCGCGGATCGCTGGAGCGGCTGTGGCAGCATAGGCTTGGG	4260
DB	3121	CTGACCCGACCTTACCATGACAGCAGCCGTCGCAACAGCGGCTGTGACGTCG	3180	QY	4261	AAGGTGCTTCTGATATTTTGGCAGGTTATGGAGCAGGCTGGCAGCGGCTGTGGGCT	4320
QY	3181	CAGCGCGAGCAGGACTGGTAGGGGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA	3240	DB	4261	AAGGTGCTTCTGATATTTTGGCAGGTTATGGAGCAGGCTGGCAGCGGCTGTGGGCT	4320
DB	3181	CAGCGCGAGCAGGACTGGTAGGGGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA	3240	QY	4321	TTTAAAGTCTAGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTTACTTCCCTGCT	4380
QY	3241	GAACGGCCCTCGGCGATGTTGATTCCTCGGTTCTGTGCGAGTCTATGACGCGGGCTGT	3300	DB	4321	TTTAAAGTCTAGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTTACTTCCCTGCT	4380
DB	3241	GAACGGCCCTCGGCGATGTTGATTCCTCGGTTCTGTGCGAGTCTATGACGCGGGCTGT	3300	QY	4381	ATCCTCTCCCTGGCGCCCTAGTCTGTGGGCTGTGTGGCAGCGATCTGCTCGGCGC	4440
QY	3301	GCTTGTACAGCTCAGCGCGCGGACCTCAGTTAGGTTGGGCTTACCTAAACACA	3360	DB	4381	ATCCTCTCCCTGGCGCCCTAGTCTGTGGGCTGTGTGGCAGCGATCTGCTCGGCGC	4440
DB	3301	GCTTGTACAGCTCAGCGCGCGGACCTCAGTTAGGTTGGGCTTACCTAAACACA	3360	QY	4441	GTGGGCCCGAGGGGAGGGGCTGTGCAGTGAACCGGCTGTAGGTTTGGCTTCCGCG	4500
QY	3361	CCAGGGTTGCCCTCTGCCAGGACCACTGGAGTTCTGGAGAGCGCTTTTACAGGGCTC	3420	DB	4441	GTGGGCCCGAGGGGAGGGGCTGTGCAGTGAACCGGCTGTAGGTTTGGCTTCCGCG	4500
DB	3361	CCAGGGTTGCCCTCTGCCAGGACCACTGGAGTTCTGGAGAGCGCTTTTACAGGGCTC	3420	QY	4501	GGTAAACAGGCTCTCCCGCCACGCACTATGTCCCTGAGAGCAGCGCTGAGCAGCTGTCAC	4560
QY	3421	ACCCATAGAGCGCCATTTCTGTGCCAGCTAAAGCAGGACAGGACAACTTCCCGCTAC	3480	DB	4501	GGTAAACAGGCTCTCCCGCCACGCACTATGTCCCTGAGAGCAGCGCTGAGCAGCTGTCAC	4560
DB	3421	ACCCATAGAGCGCCATTTCTGTGCCAGCTAAAGCAGGACAGGACAACTTCCCGCTAC	3480	QY	4561	CAGATCTCTTAGTCTTACCATCTCAGCTGCTGAAAGGCTTCCAGTGGATCAAC	4620
QY	3481	CTGTTAGCATACAGGCTACGGTGTGGCGAGGGCTCAGGCTCCACCTCATCTGTGGAC	3540	DB	4561	CAGATCTCTTAGTCTTACCATCTCAGCTGCTGAAAGGCTTCCAGTGGATCAAC	4620
DB	3481	CTGTTAGCATACAGGCTACGGTGTGGCGAGGGCTCAGGCTCCACCTCATCTGTGGAC	3540	QY	4621	GAGGATCTGCTCCACGCACTGCTCCGCTGTGGCTGAAGAGATGTTTGGGATTTGGATATGC	4680



QY	6841	ACGAGCTGCGGTAAATA	CCCTCACATGTTTACT	TGAAGCGCGCTGCGGCTCTGCAGCTGCG	6900
Db	6841	ACCAGCTGCGGTAAATAC	CCCTCACATGTTTACT	TGAAGCGCGCTGCGGCTCTGCAGCTGCG	6900
QY	6901	AAGCTCCAGGACTCAGCATG	CTCGTATGCGGAGACGACCTTGT	CGTTATCTGTGTAAGC	6960
Db	6901	AAGCTCCAGGACTCAGCATG	CTCGTATGCGGAGACGACCTTGT	CGTTATCTGTGTAAGC	6960
QY	6961	GCGGGGACCCAAAGAGCAG	AGCGCGAGCCTACGCGGCTTAC	CGGAGGCTTACGACTAGATAC	7020
Db	6961	GCGGGGACCCAAAGAGCAG	AGCGCGAGCCTACGCGGCTTAC	CGGAGGCTTACGACTAGATAC	7020
QY	7021	TCTGCCCCCTGGGACCGCC	CAACAGAAATACGACTTGGAGTTG	ATAACATCATGC	7080
Db	7021	TCTGCCCCCTGGGACCGCC	CAACAGAAATACGACTTGGAGTTG	ATAACATCATGC	7080
QY	7081	TCTTCCAAATGTTCAGTTCG	GCAGATGCTGCAAAAGGTGTACT	TCTCACCCGT	7140
Db	7081	TCTTCCAAATGTTCAGTTCG	GCAGATGCTGCAAAAGGTGTACT	TCTCACCCGT	7140
QY	7141	GACCCACCAACCCCTTGG	CGGGCTGCGTGGAGACAGCT	TACACATCAAT	7200
Db	7141	GACCCACCAACCCCTTGG	CGGGCTGCGTGGAGACAGCT	TACACATCAAT	7200
QY	7201	TCTTGGCTAGGCAACATCAT	GTATGTCGCGCCACCTTGTGGG	CAAGATGATCCTCATG	7260
Db	7201	TCTTGGCTAGGCAACATCAT	GTATGTCGCGCCACCTTGTGGG	CAAGATGATCCTCATG	7260
QY	7261	ACTCATTTCTTCCATCCCTT	CAGTTCAGGACAACTTGA	AAAGCCCTAGATTGTACG	7320
Db	7261	ACTCATTTCTTCCATCCCTT	CAGTTCAGGACAACTTGA	AAAGCCCTAGATTGTACG	7320
QY	7321	ATCTACGGGCGCTGTACT	CCATTGAGCCACTTGACCTAC	CTCAGATCAATTCACGACTC	7380
Db	7321	ATCTACGGGCGCTGTACT	CCATTGAGCCACTTGACCTAC	CTCAGATCAATTCACGACTC	7380
QY	7381	CATGGCCCTTAGCGATTTT	CACTCCATAGTTATCTCCAGG	TGAGATCAATAGGTTGGCT	7440
Db	7381	CATGGCCCTTAGCGATTTT	CACTCCATAGTTATCTCCAGG	TGAGATCAATAGGTTGGCT	7440
QY	7441	TCATGCTCAGGAAACTTGGG	GTACCGCCCTCGGAGCATCGG	GCAGACATCGGCGCAGAACT	7500
Db	7441	TCATGCTCAGGAAACTTGGG	GTACCGCCCTCGGAGCATCGG	GCAGACATCGGCGCAGAACT	7500
QY	7501	GTCGCGCTAGGCTACTGT	CCGAGGGGGAGGGCTGCCA	CTTGTGCAAGTACCTCTTC	7560
Db	7501	GTCGCGCTAGGCTACTGT	CCGAGGGGGAGGGCTGCCA	CTTGTGCAAGTACCTCTTC	7560
QY	7561	AACTGGGCAGTAGGACCAAG	CTCAACTCATCTCAATCCCGG	TGCGTCCCAAGTTGGAT	7620
Db	7561	AACTGGGCAGTAGGACCAAG	CTCAACTCATCTCAATCCCGG	TGCGTCCCAAGTTGGAT	7620
QY	7621	TTATCAGCTGGTTCTGTCTG	TTACAGGGGGAGACATATAC	CACGCTGTCTCGT	7680
Db	7621	TTATCAGCTGGTTCTGTCTG	TTACAGGGGGAGACATATAC	CACGCTGTCTCGT	7680
QY	7681	GCCGACCCCGCTGGTTCA	TGTGCTTCTGTAGGGGTAGG	CACTCTAT	7740
Db	7681	GCCGACCCCGCTGGTTCA	TGTGCTTCTGTAGGGGTAGG	CACTCTAT	7740
QY	7741	CTACTCCCAACCGATGAAC	GGGACCTAAACCTCCAGGCC	CAATAGGCCATCCTGTTT	7800
Db	7741	CTACTCCCAACCGATGAAC	GGGAGCTAAACCTCCAGGCC	CAATAGGCCATCCTGTTT	7800
QY	7801	TTTCCCTTTTTTTTTCTTTT	TTTTTTTTTTTTTTTTTTTTTT	TTTTTTTTTTTTTTTTTTTTTT	7860
Db	7801	TTTCCCTTTTTTTTTCTTTT	TTTTTTTTTTTTTTTTTTTTTT	TTTTTTTTTTTTTTTTTTTTTT	7860
QY	7861	TTTTTCTCTTTTTTTTCTTTT	TTTTTTTTTTTTTTTTTTTTTT	TTTTTTTTTTTTTTTTTTTTTT	7920
Db	7861	TTTTTCTCTTTTTTTTCTTTT	TTTTTTTTTTTTTTTTTTTTTT	TTTTTTTTTTTTTTTTTTTTTT	7920
QY	7921	TAGCTGTGAAGGTCGCTGAG	CGCGCTGTACTGACGAGAGT	GTGTGATCTAGGCTCTCTGC	7980

Db	7921	TAGCTGTGAAGGTCCTGTAGCGCGCTTGACTGCAGAGAGTGCTGATACTGGCCTCTCTGC	7980
Qy	7981	AGATCAAGT	7989
Db	7981	AGATCAAGT	7989
RESULT 2			
LOCUS	AX036255		
DEFINITION	Sequence 4 from Patent EP1043399.	7989 bp	DNA
ACCESSION	AX036255		linear
VERSION	AX036255.1	GI:11225871	
KEYWORDS			
SOURCE	Hepatitis C virus		
ORGANISM	Hepatitis C virus		
REFERENCE	1		
AUTHORS	Bartschlagher,R.D.		
TITLE	Hepatitis C virus cell culture system		
JOURNAL	Patent: EP 1043399-A 4 11-OCT-2000;		
FEATURES	BARTSCHLAGHER RALF DR (DE)		
source	Location/Qualifiers		
	1..7989		
	/organism="Hepatitis C virus"		
	/mol_type="unassigned DNA"		
ORIGIN	/db_xref="taxon:11103"		
Query Match	100.0%;	Score 7987.4;	DB 6; Length 7989;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 7989;	Conservative 0;	Mismatches 1;	Indels 0; Gaps 0;
Qy	1	GCCAGCCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTCTGTGAGGAATCTACTG	60
Db	1	GCCAGCCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTCTGTGAGGAATCTACTG	60
Qy	61	TCTTACGACGAGAAGGCTTAGCCATGGCGTTAGTATGAGTGTCTGTCAGCCTCCAGGAC	120
Db	61	TCTTACGACGAGAAGGCTTAGCCATGGCGTTAGTATGAGTGTCTGTCAGCCTCCAGGAC	120
Qy	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG	180
Db	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG	180
Qy	181	GAGACCGGGTCTTTCTTTGGATCAACCCGCTCAATGCTCGAGATTTGGGCGTGCCTCC	240
Db	181	GAGACCGGGTCTTTCTTTGGATCAACCCGCTCAATGCTCGAGATTTGGGCGTGCCTCC	240
Qy	241	GCGAGACTCTAGCCAGTAGTGTGGGTGCGGAAGGCTTTGGTACTGCCTGATAGG	300
Db	241	GCGAGACTCTAGCCAGTAGTGTGGGTGCGGAAGGCTTTGGTACTGCCTGATAGG	300
Qy	301	GTGCTTGAGTGCCTCCCGGAGGCTCTCGTAGACCGTGCACCATAGACGAGATCTTAAC	360
Db	301	GTGCTTGAGTGCCTCCCGGAGGCTCTCGTAGACCGTGCACCATAGACGAGATCTTAAC	360
Qy	361	CTCAAGAAAAACCAAGGCGCGGCATGATTGAACAAGATGGATTGCACGCGAGTTCTC	420
Db	361	CTCAAGAAAAACCAAGGCGCGGCATGATTGAACAAGATGGATTGCACGCGAGTTCTC	420
Qy	421	CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAACAGCAATCGCTGCT	480
Db	421	CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAACAGCAATCGCTGCT	480
Qy	481	CTGATGCCGCGTGTTCGGCTGTTCAGCGAGGGGCGCGGTTCTTTTGTCAAGACCG	540
Db	481	CTGATGCCGCGTGTTCGGCTGTTCAGCGAGGGGCGCGGTTCTTTTGTCAAGACCG	540
Qy	541	ACCTCTGCGGTCCTGTAATGAATGACGAGGACGAGGCGCGGCTATCTGCTGGCTGGCCA	600



QY	2761	CTGCGCACCGCTACGCTCGGGATCGGTCAACCGTCCCATCATCAACATCGAGAGGTG	2820
Db	2761		2820
QY	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTATATGGCAAGCCATCCCCATCGAGACCATC	2880
Db	2821		2880
QY	2881	AAGGGGGAGGACCTCATTTTCTGCCATTCCAAAGAAATATGTATAGCTCGCGCG	2940
Db	2881		2940
QY	2941	AAGCTGTCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATGATCCGTC	3000
Db	2941		3000
QY	3001	ATACCAACTAGCGAGACGTCATTGTGTAGCAACGAGACGCTCTAATGACGGGCTTTACC	3060
Db	3001		3060
QY	3061	GGCGATTTGACTCGATCGACTGCAATACATATATGTGTACCCAGACAGTCGACTTCAGC	3120
Db	3061		3120
QY	3121	CTGGACCGGACCTTCACTTGTAGAGACGACCGTGCACAGACGGGTGTCAACGCTCG	3180
Db	3121		3180
QY	3181	CAGCGCGAGGACGAGTGTGTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGGA	3240
Db	3181		3240
QY	3241	GAACGGCCCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGTGT	3300
Db	3241		3300
QY	3301	GCTTGTGACAGCTACGCGCGCGGACCTCAGTTAGGTTGCGGCTTACCTAAACACA	3360
Db	3301		3360
QY	3361	CCAGGTTGCCGTCTGCCAGGACCATCTGGAGTTCTGGAGAGCGTCTTTACAGGCTC	3420
Db	3361		3420
QY	3421	ACCACATAGACGCCATTTCTTGTCCAGACTAAGCAGCAGGAGACAACTTCCCGTAC	3480
Db	3421		3480
QY	3481	CTGGTAGCATACAGGCTACGGTGTGGCGCAGGGCTCAGGCTCCACCTCATCGTGGAC	3540
Db	3481		3540
QY	3541	CAAAATGTGAAGTGTCTCATACGGCTAAAGGCTACGCTGACGCGGCCCAACGCCCTGCTG	3600
Db	3541		3600
QY	3601	TATAGGCTGGAGCGGTTCAAAACGAGGTTTACTACCAACACACCCCATTAACCAATATATC	3660
Db	3601		3660
QY	3661	ATGGCATGATGTGCGGTGAGGTCGTGTACAGGACCTGGGTGCTGTAGGCGGA	3720
Db	3661		3720
QY	3721	GTCTTAGCAGCTGTGGCGCGGTATTTGCTGACAAACAGGACGCTGTGCTATTGTGGGCGAG	3780
Db	3721		3780
QY	3781	ATCATTTTGTCCGAAGCCGGCATCATTTCCGACAGGAAAGTCCTTTTACCGGAGTTC	3840
Db	3781		3840
QY	3841	GATGAGATGGAAGTGGGCTCACACCTCCCTTACATCGAACAGGGAATGACGCTCGCC	3900
Db	3841		3900

Db	3841	GATGAGATGGAAGTGGGCTCACACCTCCCTTACATCGAACAGGGAATGACGCTCGCC	3900
QY	3901	GAACAATTTCAACAGAGCAATCGGGTTCCTGCAACAGCCACCAAGCAAGCGGAGGCT	3960
Db	3901		3960
QY	3961	GCTGCTCCCGTGTGGAATCCAAGTGGCGACCTCGAAGCCCTTCTGGCGAAGCATATG	4020
Db	3961		4020
QY	4021	TGCAATTTTCAATCAGCGGATACAAATTTAGCAGGCTTGTCCACTGTGCTGCGCAAGCC	4080
Db	4021		4080
QY	4081	GCGATAGCATCACTGATGGCATTCACAGCCCTTATCACAGCCCGCTCACACCCCAACAT	4140
Db	4081		4140
QY	4141	ACCTCTCTGTTTAAATCTCTGTGGGGGATGGTGGCGGCCCACTTGTCTCTCCAGGCT	4200
Db	4141		4200
QY	4201	GCTTCTGCTTTCGTAGCGCGGCATCGCTGGAGCGGCTTGTGCGACATAGGCTTGGG	4260
Db	4201		4260
QY	4261	AAGGTGCTTGTGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCGCTCGTGCC	4320
Db	4261		4320
QY	4321	TTTAAAGTCTATGAGCGCGGATGCCCCCTCACCGAGGACCTGGTTAACTTCTCCCTGT	4380
Db	4321		4380
QY	4381	ATCCTCTCCCTCGCGCCCTAGTGTGGGGTGTGTGCGCAGCGATCTGCGCTGGCAC	4440
Db	4381		4440
QY	4441	GTGGGCCCGAGGGGCGGTGTGCAGTGGATGAACCGGCTGATAGCGTTTCGCTTCGCGG	4500
Db	4441		4500
QY	4501	GGTAAACAGCTTCCCGCCAGCACTATGCTGTGAGAGGACGCTGAGCAGCTGTGCTAC	4560
Db	4501		4560
QY	4561	CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGAGAGGCTTCCACGATGGATCAAC	4620
Db	4561		4620
QY	4621	GAGGACTGTCTCACGCCATGCTCCGGCTGTGGCTAAGAGATGTTTGGGATTTGATATGC	4680
Db	4621		4680
QY	4681	ACGGTGTGTGATTTTCAAGACCTGGCTCAGTCCAAGCTCTCTGCGCGGATTTGCGGGA	4740
Db	4681		4740
QY	4741	GTCCCTCTCTCTCATGTCAACGTGGGTACAGGGAGTCTGGCGGGCGACGCGCATATG	4800
Db	4741		4800
QY	4801	CAAAACCACTGCCATGTGGAGCACAGATCACCGGACATGTGAAAAACGGTTCCATGAGG	4860
Db	4801		4860
QY	4861	ATCGTGGGCTTAGGACCTGTAGTAAACGCTGATGGAACATTTCCCAATTAAACGGGTAC	4920
Db	4861		4920
QY	4921	ACCAACGGGCTTGCACGCTCCCGGCCCAATTTATCTAGGGGCTGTGGCGGGTG	4980
Db	4921		4980

Db 4921 ACCACGGGCCCTGCAAGCCCTCCCGCGGCCCAAAATTATTCTAGGCGCTGTGGCGGTG 4980  
Qy 4981 GCTGCTGAGGAGTAGTGGAGGTTACGCGGGTGGGGGATTTCCACTACGTGACGGGCATG 5040  
Db 4981 GCTGCTGAGGAGTAGTGGAGGTTACGCGGGTGGGGGATTTCCACTACGTGACGGGCATG 5040  
Qy 5041 ACCACTGACAAAGTAAAGTGCCTGTCAAGTTTCAGGTTCCGGCCCCCGAAATTTCTACAGAAGTG 5100  
Db 5041 ACCACTGACAAAGTAAAGTGCCTGTCAAGTTTCAGGTTTCGGCCCCCGAAATTTCTACAGAAGTG 5100  
Qy 5101 GATGGGTGCGGTTCGACAGGTACGCTCCAGCGTGCAGAACCCCTCCTACGGGAGGAGGTC 5160  
Db 5101 GATGGGTGCGGTTCGACAGGTACGCTCCAGCGTGCAGAACCCCTCCTACGGGAGGAGGTC 5160  
Qy 5161 ACATTCCTGGTGGGCTCAATCAATACCTGTTGGTTCAGGTTCAGGTTTCGGCCCCCGAAATTTCTACAGAAGTG 5220  
Db 5161 ACATTCCTGGTGGGCTCAATCAATACCTGTTGGTTCAGGTTCAGGTTTCGGCCCCCGAAATTTCTACAGAAGTG 5220  
Qy 5221 CCGGACGTAGCAGTGTCACTTCCATGCTCACCGACCCCTCCCAATACGGCGGAGACG 5280  
Db 5221 CCGGACGTAGCAGTGTCACTTCCATGCTCACCGACCCCTCCCAATACGGCGGAGACG 5280  
Qy 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTGGCCAGCTCATCAGCTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTGGCCAGCTCATCAGCTAGCCAG 5340  
Qy 5341 CTGCTCTGGCCCTTCTTGAAGCAACATGCTCACCGACCCCTCCCAATACGGCGGAGACG 5400  
Db 5341 CTGCTCTGGCCCTTCTTGAAGCAACATGCTCACCGACCCCTCCCAATACGGCGGAGACG 5400  
Qy 5401 CTGCTCTGGCCCTTCTTGAAGCAACATGCTCACCGACCCCTCCCAATACGGCGGAGACG 5460  
Db 5401 CTGCTCTGGCCCTTCTTGAAGCAACATGCTCACCGACCCCTCCCAATACGGCGGAGACG 5460  
Qy 5461 TCGAAAAATPAGGTAGTAATTTGGACTCTTTGGAGCTCTTTCGAGCCCTCCAAAGCGGAGGAGTAG 5520  
Db 5461 TCGAAAAATPAGGTAGTAATTTGGACTCTTTGGAGCTCTTTCGAGCCCTCCAAAGCGGAGGAGTAG 5520  
Qy 5521 AGGGAAGTATCCGTTCCGGCGGAGATCTCGGAGGTTCAGGAATTCCTCGAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCGGCGGAGATCTCGGAGGTTCAGGAATTCCTCGAGCGATG 5580  
Qy 5581 CCCATATGGCGACCGCCGATTTACAACTTCACTGTTAGAGTCTGGAAGGACCCCGGAC 5640  
Db 5581 CCCATATGGCGACCGCCGATTTACAACTTCACTGTTAGAGTCTGGAAGGACCCCGGAC 5640  
Qy 5641 TACGTCCCTCCAGTGTGTAACCGGTGTCATTCGCGCTGCAAGGCGCTCCGATACCA 5700  
Db 5641 TACGTCCCTCCAGTGTGTAACCGGTGTCATTCGCGCTGCAAGGCGCTCCGATACCA 5700  
Qy 5701 CCTCCACGAGGAGGAGGAGGTTGCTGTCAGATCTACCGTGTCTTCTGCTTGGCG 5760  
Db 5701 CCTCCACGAGGAGGAGGAGGTTGCTGTCAGATCTACCGTGTCTTCTGCTTGGCG 5760  
Qy 5761 GAGCTCGCCACAAAGACCTTCGCGAGCTCCGAAATCGTCGGCCGTGCAGACGGCGACGGCA 5820  
Db 5761 GAGCTCGCCACAAAGACCTTCGCGAGCTCCGAAATCGTCGGCCGTGCAGACGGCGACGGCA 5820  
Qy 5821 ACGGCTCTCTGACACAGCCCTCCGACGACGGCGGATCCGAACTTCGAGTCGTAC 5880  
Db 5821 ACGGCTCTCTGACACAGCCCTCCGACGACGGCGGATCCGAACTTCGAGTCGTAC 5880  
Qy 5881 TCCTCCATGCCCCCTTCAGGGGAGCCGGGATCCCGATCTCAGACGGGCTTGG 5940  
Db 5881 TCCTCCATGCCCCCTTCAGGGGAGCCGGGATCCCGATCTCAGACGGGCTTGG 5940  
Qy 5941 TCTACCGTAAGGAGGAGGCTAGTGAGGACGTCGTCTGCTGCTCGATGTCCTACACATGG 6000  
Db 5941 TCTACCGTAAGGAGGAGGCTAGTGAGGACGTCGTCTGCTGCTCGATGTCCTACACATGG 6000  
Qy 6001 ACAGGCGCCCTGATACCGCCATGCGTGGGAGGAAACCAAGTGCCTCCATATGACATG 6060  
Db 6001 ACAGGCGCCCTGATACCGCCATGCGTGGGAGGAAACCAAGTGCCTCCATATGACATG 6060

Qy 6061 AGCAACTCTTTGCTCCGTCAACCAAACTTTGGTCTATGCTACACATCTCGACGCGCAAGC 6120  
Db 6061 AGCAACTCTTTGCTCCGTCAACCAAACTTTGGTCTATGCTACACATCTCGACGCGCAAGC 6120  
Qy 6121 CTGCGCGCAGAAAGGTCACTTTTCAGACACTGCAAGTCTCTGAGCGACCACTACCCGGAC 6180  
Db 6121 CTGCGCGCAGAAAGGTCACTTTTCAGACACTGCAAGTCTCTGAGCGACCACTACCCGGAC 6180  
Qy 6181 GTGCTCAAGAGATGAAGCGAAGCGTCCACAGTTAAGGCTAAACTTTATCCGTGGAG 6240  
Db 6181 GTGCTCAAGAGATGAAGCGAAGCGTCCACAGTTAAGGCTAAACTTTATCCGTGGAG 6240  
Qy 6241 GAAGCCTGTAGCTGACGCCCCCAATTCGGCCAGATCTAAATTTGGCTATCGGCAAG 6300  
Db 6241 GAAGCCTGTAGCTGACGCCCCCAATTCGGCCAGATCTAAATTTGGCTATCGGCAAG 6300  
Qy 6301 GACGTCGCGAACCTATCCAGCAAGCGGTTAACCACTCCGCTCCGTGTGGAAGGACTTG 6360  
Db 6301 GACGTCGCGAACCTATCCAGCAAGCGGTTAACCACTCCGCTCCGTGTGGAAGGACTTG 6360  
Qy 6361 CTGGAAGACACTGAGACACCAATTCGACCACTCATGCGCAAAAAATGAGGTTTTCTGC 6420  
Db 6361 CTGGAAGACACTGAGACACCAATTCGACCACTCATGCGCAAAAAATGAGGTTTTCTGC 6420  
Qy 6421 GTCCAAACAGAGAAAGGGGGCGCAAGCGAGCTCGCCCTATCGTATTCACGATTTGGG 6480  
Db 6421 GTCCAAACAGAGAAAGGGGGCGCAAGCGAGCTCGCCCTATCGTATTCACGATTTGGG 6480  
Qy 6481 GTTCGTGTGTGCGAGAAATGGCCCTTTACAGTGTGTCTCCACCTCCCTCAGGCGGTG 6540  
Db 6481 GTTCGTGTGTGCGAGAAATGGCCCTTTACAGTGTGTCTCCACCTCCCTCAGGCGGTG 6540  
Qy 6541 ATGGGCTCTTCAATACGAGTTCCAAATCTCTCTGACACGCGGTGAGTTCCTGTGAAAT 6600  
Db 6541 ATGGGCTCTTCAATACGAGTTCCAAATCTCTCTGACACGCGGTGAGTTCCTGTGAAAT 6600  
Qy 6601 GCTCGGAAGCGAAGAAATGGCTATGGCTTCGCAATGACACCGCTGTTTTCAGTCA 6660  
Db 6601 GCTCGGAAGCGAAGAAATGGCTATGGCTTCGCAATGACACCGCTGTTTTCAGTCA 6660  
Qy 6661 ACAGTCACTGAGAAATGACATCCGTGTGAGGAGTCAATCTACCAATGTTGTGACTGGCC 6720  
Db 6661 ACAGTCACTGAGAAATGACATCCGTGTGAGGAGTCAATCTACCAATGTTGTGACTGGCC 6720  
Qy 6721 CCGGAAGCGACAGCGCATAGGTGCTTCAGACGCGCTTTACATCGGGGGCCCCCTG 6780  
Db 6721 CCGGAAGCGACAGCGCATAGGTGCTTCAGACGCGCTTTACATCGGGGGCCCCCTG 6780  
Qy 6781 ACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGAGCGGTGACTGACG 6840  
Db 6781 ACTAATTTCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGAGCGGTGACTGACG 6840  
Qy 6841 ACCAGTCCGGTAAATACCCCTACATGTTTACCTTGAAGGCGCTGCGGCTGTGAGTGGC 6900  
Db 6841 ACCAGTCCGGTAAATACCCCTACATGTTTACCTTGAAGGCGCTGCGGCTGTGAGTGGC 6900  
Qy 6901 AAGCTCCAGGACTGACAGATGCTGTAAGCGAGACGCTTGTGTTATCTGTGAAAGC 6960  
Db 6901 AAGCTCCAGGACTGACAGATGCTGTAAGCGAGACGCTTGTGTTATCTGTGAAAGC 6960  
Qy 6961 GCGGGGACCCAAAGAGAGAGCGAGCGCTTACGGGCTTACAGAGGCTATGACTAGATAC 7020  
Db 6961 GCGGGGACCCAAAGAGAGAGCGAGCGCTTACGGGCTTACAGAGGCTATGACTAGATAC 7020  
Qy 7021 TCTGCCCCCTTGGGACCCGCCAAACAGAAATACGACTTGGAGTTGATAACATCATGC 7080  
Db 7021 TCTGCCCCCTTGGGACCCGCCAAACAGAAATACGACTTGGAGTTGATAACATCATGC 7080  
Qy 7081 TCTTCAATGTGTGAGTCCGCGACGATGCTGCGCAAGGAGGTTGACTATCTCACCCGT 7140  
Db 7081 TCTTCAATGTGTGAGTCCGCGACGATGCTGCGCAAGGAGGTTGACTATCTCACCCGT 7140



```
QY 7141 GACCCACACACCCCTTGGCGGGCTGGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
|||
Db 7141 GACCCACACACCCCTTGGCGGGCTGGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
|||
QY 7201 TCTGCTAGGCAACATCATCATGTATGCGCCACCTTGGGCAAGGATGATCTCTGATG 7260
|||
Db 7201 TCTGCTAGGCAACATCATCATGTATGCGCCACCTTGGGCAAGGATGATCTCTGATG 7260
|||
QY 7261 ACTCATTTCTTCCATCTCTTCTAGCTCAGGAAACAACTGAAAAAGCCCTTAGATTCTCAG 7320
|||
Db 7261 ACTCATTTCTTCCATCTCTTCTAGCTCAGGAAACAACTGAAAAAGCCCTTAGATTCTCAG 7320
|||
QY 7321 ATCTAGGGCGCTTACTCCATTTAGGCACTTGCCTCAGCTCAGATCATCTCAACGACTC 7380
|||
Db 7321 ATCTAGGGCGCTTACTCCATTTAGGCACTTGCCTCAGCTCAGATCATCTCAACGACTC 7380
|||
QY 7381 CATGGCTTAGCGCATTTTCACTCCATAGTATCTCTCCAGGTGAGATCAATAGGTGGCT 7440
|||
Db 7381 CATGGCTTAGCGCATTTTCACTCCATAGTATCTCTCCAGGTGAGATCAATAGGTGGCT 7440
|||
QY 7441 TCATGCTCAGGAAACTTGGGGTACCGCCCTTGCAGTCTGGAGACATCGGGCCAGAGT 7500
|||
Db 7441 TCATGCTCAGGAAACTTGGGGTACCGCCCTTGCAGTCTGGAGACATCGGGCCAGAGT 7500
|||
QY 7501 GTCGGCTAGGCTACTGTCACAGGGGGAGGCTGCCACTTGTGCAAGTACCTCTTC 7560
|||
Db 7501 GTCGGCTAGGCTACTGTCACAGGGGGAGGCTGCCACTTGTGCAAGTACCTCTTC 7560
|||
QY 7561 AACTGGGCACTAAGGACCAAGCTCAAACTCACTCCAACTCCCGCTGGTCCCAAGTTGGAT 7620
|||
Db 7561 AACTGGGCACTAAGGACCAAGCTCAAACTCACTCCAACTCCCGCTGGTCCCAAGTTGGAT 7620
|||
QY 7621 TTATCCAGCTGGTTCGTGCTGTACAGGGGGAGACATATATACAGCCCTGTCTCGT 7680
|||
Db 7621 TTATCCAGCTGGTTCGTGCTGTACAGGGGGAGACATATATACAGCCCTGTCTCGT 7680
|||
QY 7681 GCCGACCCCGCTGTTCTGCTGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7740
|||
Db 7681 GCCGACCCCGCTGTTCTGCTGCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 7740
|||
QY 7741 CTACTCCCAACCGATGAACGGGACCTAACTCACTCCAGGCCAATAGGCCATCTCTGTTT 7800
|||
Db 7741 CTACTCCCAACCGATGAACGGGACCTAACTCACTCCAGGCCAATAGGCCATCTCTGTTT 7800
|||
QY 7801 TTTCCCTTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7860
|||
Db 7801 TTTCCCTTTTCTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7860
|||
QY 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
|||
Db 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
|||
QY 7921 TAGCTGTGAAGCTCCGCTGAGCGCTTGACTGCAGAGAGTGTCTGATCTGGCCTCTCTGC 7980
|||
Db 7921 TAGCTGTGAAGCTCCGCTGAGCGCTTGACTGCAGAGAGTGTCTGATCTGGCCTCTCTGC 7980
|||
QY 7981 AGATCAAGT 7989
|||
Db 7981 AGATCAAGT 7989
|||
```

```
RESULT 3
SSE242652
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SSE242652
Hepatitis C virus replicon 1377/NS3-3'UTR.
AJ242652
AJ242652.1 GI:5441834
core-neo gene; NS2 gene; NS2 proteinase; NS3 gene; NS3
proteinase/helicase; NS3/4A proteinase cofactor; NS4A gene; NS4B
gene; NS4B protein; NS5A gene; NS5A phosphoprotein; NS5B gene; NS5B
RNA dependant RNA polymerase; polyprotein.
Hepatitis C virus replicon 1377/NS3-3'UTR
Hepatitis C virus replicon 1377/NS3-3'UTR
```

```
artificial sequences; vectors.
1
Lohmann,V., Korner,F., Koch,J., Herian,U., Theilmann,L. and
Bartenschlager,R.
Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell
line
Science 285 (5424), 110-113 (1999)
99322193
MEDLINE
10390360
PUBMED
2 (bases 1 to 7989)
Bartenschlager,R.
Direct Submission
Submitted (26-MAY-1999) Bartenschlager R., Institute for Virology,
Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,
55131 Mainz, GERMANY
FEATURES
Location/Qualifiers
1..7989
/organism="Hepatitis C virus replicon I377/NS3-3'UTR"
/mol_type="other RNA"
/db_xref="taxon:95363"
/focus
/notes="Neomycin-selectable bicistronic subgenomic
hepatitis C virus replicon"
join(1..341,1801..7758)
/organism="Hepatitis C virus"
/mol_type="other RNA"
/strain="fb"
/isolates="Conl"
/db_xref="taxon:11103"
/tissue_type="liver"
1190..1800
/organism="Encephalomyocarditis virus"
/mol_type="other RNA"
/db_xref="taxon:12104"
1..341
/evidence="experimental"
342..1181
/genes="core-neo"
342..1181
/function="selectable marker"
/codon_start=1
/product="core-neo fusion protein"
/protein_id="CAB46912.1"
/db_xref="GI:5441835"
/translation="MSTNPKPQKTKGRAMIEODGLHAGSPAAWVERLFQYDWAQOTTI
GSDAAVFLSAQREVLVKTLDSGALNELQDEAARLSWLAATGVPQARSLTPTCTGCS
GRDWLLGLGVPGQDLLSHLAPAEKVSINADAMRRLHLDLPATCPDPCAARLIERAR
TRMEAGLVDDDLDEHQGLAPAEKPARLKAMPDGEDLVVTHGDACLPMIVENGRF
SGFDICGLGVADRYQDIALATRDIAELGGEWADRFVLVLTAAFDSPQRIAFYRLLD
EPF"
misc_feature
1190..1800
/notes="internal ribosome entry site (IRES)"
1801..7758
/codon_start=1
/product="non-structural polyprotein"
/protein_id="CAB46913.1"
/db_xref="GI:5441836"
/translation="MAPITAYSQQTGRLGCIITSLTGRDRNQVEGEVQVSTATQSOF
LATCVNGCVTVVYVAGSKTLAGPKGPIQMTVNTVDLGVQAPGASRLTPTCTGCS
SDLYLTHRADVI PVERRGDSRGLLSPREVSLKSGSGPLLCPSGHAGVIFRAAVC
TRGVAKAVDFPVESNETMRSPTDSSPPAVPOTFOVAHLHAPTGSCKSTKVPA
YAAQGGKVDVLPNSVAATLGFAYMSKAGHDINIKITGTTGAPITYSTYKGLA
DGGCSGKYVDIIICDECHSTDSITLIGITVLDDQAEAGARLVLTATPPGVSIVPH
PNIIEVALSGTBIPEYKAIPIETIKGRLHIFCHSKKCDLAKLGLGLNVAAY
YRGLDVSIVPTGDIIVATDALMTGTFGDSVITDNTCVTQVDFSLDPTIETT
TVPDQAVSRQRGRTGRGMGIYRPVTPERSGEMFSDSVLCECYDAGCAWELTPA
TVSRRLAYINTPGLPVCDHLEFFWESVETGLTHIDAPLSQTKQAGDNFFVLVQA
TVCRAQAPPPSWDQMKLIRLPLRHGTPLLYRLGAVQNEVTHIPIITIMACM
SADLEVVTWVLVGGVLAALAAAYCLTTGSGVIVGRIILSGKPAIIPDREVLRFED
MECSHLPIEQGMQLAFQKAKGLLQTKAQEAAPVSVESKWRLEAFWAKHM
WNFISGIOVLAGLSTLPNPAJASLMAFTASITSLPTQHTLLFNILGWWAAQLAPP
SAASAFVGAGIAGAAVSGISGLGKLVLDILAGYGAGVAGALVAFKVMSEMPSTEDLVN
```

LLPALLSPALVGVVCAAILRRHVGPGEVQVMNRLIAFASRGNHVSPHYPESD  
AAARVTOILLSSLTITQLKRLHOWINEDCSTPCSGWLRDWDMLCTVLTOPKTLQOS  
KLLPLRPGVPFSCQGVYGRGIMQITTCPCQAQITGHVKNCSMLRGVPTCSNT  
WHGTPEINAYTGPCTPAPNYSRALMRVAABEYEVTRVGDPRHYVTGMTIDNVKCP  
CQVPAPEFTEVDGVRLLHRYAPACPLUREEVTFLVNLQYLVSQLPCEPDPVAVL  
TSMLEDPSHI TAETAKRLARSPSLASSSQLSAPSLKATCTRHDSPPADLIEA  
NLLMRQMGNTIRVSESNKVILDSFELPQAEDEREVSPAEILRRSRKPPRAMPI  
WARPDYNPLLESDDPDVVPVHVHCPPLPAKAPPIPPRRKRTVVLSESTVSALA  
ELATVTFSSSESSAVDSGTATASDPDSDGDGSDVESYSSMPPLEGEPPDPLSDG  
SWSVTSBEASEVDVCCSGSYWTGALIIPCAABETKLPINALNSLLRHNLVJYATTS  
RSASUKQVTFDRLOVDDHVDVLMKEMKAKASTVAKLLSVEEACKLTPPHSARSK  
FYGADKVRNCSKAVNHRSVKOLLEDTETPIDTIMAKNEVFCVOPKGRKPAR  
LIVFDPLGVNCEKMALYDVSTLPQAVMGSSYGFQSPQORVELVNAWKAKCPMG  
FAYDPRCDSTVTENDIIVEESIYOCCLAPAROAIRSLTERLVIGGLPTNSKQNC  
GYRRCASGLVLTSCGNTLTCTYLKAAACRAKLDQCTMLVCGDLDLVICBSAGTQED  
EASLRAFTEAMTRYISAPDPPKPEYDLELITSNSNVSAHDAGKRVLYLITRPTT  
PLARAWEARTHTPVNSWIGNIIMYAPTLAMILMTHFFSILLAOEOLKALDCQIY  
GACYSIEFLDLPQIIFORLHGLSAFSLHSYSGEINRVASCLFKLGVPLRVVRRHARS  
VRALLSQGGRAATCGKYLFWAVRTKLKLTPIPAASQLDLSWFEVAGVSGGDIYHSL  
SRARPRMFWMLLLLSVGVGIYLLPNR"

mat\_peptide

1801..3696

/gene="NS3"  
/product="NS3 proteinase/helicase"

mat\_peptide

3697..3858

/gene="NS4"  
/product="NS3/4A proteinase cofactor"

mat\_peptide

3859..4641

/gene="NS4B"

mat\_peptide

4642..5982

/product="NS4B protein"

mat\_peptide

5983..7755

/gene="NS5A"  
/product="NS5A phosphoprotein"

mat\_peptide

7756..7989

/gene="NS5B"  
/product="NS5B RNA dependant RNA polymerase"

gene

1801..3696

gene

3697..3858

gene

3859..4641

gene

4642..5982

gene

5983..7755

gene

7756..7989

3'UTR

ORIGIN

Query Match 100.0%; Score 7987.4; DB 12; Length 7989;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATGGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAATCTGTG 60

Db 1 GCCAGCCCCGATGGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAATCTGTG 60

QY 61 TCTTCACGACAGAAAGCGTCTAGCCATGCGTTAGTATGATGAGTGTGCGAGCTCCAGGAC 120

Db 61 TCTTCACGACAGAAAGCGTCTAGCCATGCGTTAGTATGATGAGTGTGCGAGCTCCAGGAC 120

QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATGCCAG 180

Db 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATGCCAG 180

QY 181 GACGACCGGGTCTCTTCTTGGATCAACCGCTCAATGCTGAGATTGGGCGTCCCCC 240

Db 181 GACGACCGGGTCTCTTCTTGGATCAACCGCTCAATGCTGAGATTGGGCGTCCCCC 240

QY 241 CGGAGACTGCTAGCCGAGTGTGGGTGCGAAAGGCCCTTGTGGTACTGCTGATAGG 300

Db 241 CGGAGACTGCTAGCCGAGTGTGGGTGCGAAAGGCCCTTGTGGTACTGCTGATAGG 300

QY 301 GTGCTTGAGTGGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGGAATCTAAAC 360

Db 301 GTGCTTGAGTGGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGGAATCTAAAC 360

Db 301 GTGCTTGAGTGGCCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGGAATCTAAAC 360

QY 361 CTCAAAGAAAACCAAAGAGGCGGCATGATTGAACAAGATGGATTGCACGAGGTTCTC 420

Db 361 CTCAAAGAAAACCAAAGAGGCGGCATGATTGAACAAGATGGATTGCACGAGGTTCTC 420

QY 421 CGGCGCGTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGCAATCGGCTGCT 480

Db 421 CGGCGCGTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGCAATCGGCTGCT 480

QY 481 CTGATGCGCGCGTGTTCGGGCTGTTCAGCGGACGGGCGCGCGGTTCTTTTGTCAAGACCG 540

Db 481 CTGATGCGCGCGTGTTCGGGCTGTTCAGCGGACGGGCGCGCGGTTCTTTTGTCAAGACCG 540

QY 541 ACCTGTCGGTGCCTTGAATGAATGCGAGGACGAGGCGCGGCTATCGTGGCTGGCCA 600

Db 541 ACCTGTCGGTGCCTTGAATGAATGCGAGGACGAGGCGCGGCTATCGTGGCTGGCCA 600

QY 601 CGACGGGCGTTCCTTTCGCGAGCTGTGCTGCGAGTTGTCACTGAAGCGGAAGGACTGCG 660

Db 601 CGACGGGCGTTCCTTTCGCGAGCTGTGCTGCGAGTTGTCACTGAAGCGGAAGGACTGCG 660

QY 661 TGCTATTGGGCGAAGTGCAGGCGGCGAGGATCTCCTGTCTCATCTCACCTTGTCTCTGCGAGA 720

Db 661 TGCTATTGGGCGAAGTGCAGGCGGCGAGGATCTCCTGTCTCATCTCACCTTGTCTCTGCGAGA 720

QY 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCTGATCGGCTGATCGGCTGCTGCG 780

Db 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCTGATCGGCTGATCGGCTGCTGCG 780

QY 781 CATTCGACCAACCAAGCGAACAATCGATCGAGCGAGCAGTACTCGGATGGAAGCGGTC 840

Db 781 CATTCGACCAACCAAGCGAACAATCGATCGAGCGAGCAGTACTCGGATGGAAGCGGTC 840

QY 841 TTGTCGATCAGGATGATCTGACGAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTCG 900

Db 841 TTGTCGATCAGGATGATCTGACGAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTCG 900

QY 901 CAGGCTCAAGGCGCGCATGCGCGAGGATCTCGTGTGACCAATGCGATGCGATGCGCT 960

Db 901 CAGGCTCAAGGCGCGCATGCGCGAGGATCTCGTGTGACCAATGCGATGCGATGCGCT 960

QY 961 GCTTCCGCAATATCATGTGGAAATGCGCGTTCCTGAGTATCATGACATGTCGCGCGC 1020

Db 961 GCTTCCGCAATATCATGTGGAAATGCGCGTTCCTGAGTATCATGACATGTCGCGCGC 1020

QY 1021 TGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGC 1080

Db 1021 TGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGC 1080

QY 1081 TTGGCGGGAATGGGCTGACCGCTTCCTGCTGCTTTAGGTTATCGCGCTCCCGATTGCG 1140

Db 1081 TTGGCGGGAATGGGCTGACCGCTTCCTGCTGCTTTAGGTTATCGCGCTCCCGATTGCG 1140

QY 1141 AGCGCATCGCTCTATCGCCCTCTTTCGAGAGTCTTCTGAGTATTAACAGACCAACAG 1200

Db 1141 AGCGCATCGCTCTATCGCCCTCTTTCGAGAGTCTTCTGAGTATTAACAGACCAACAG 1200

QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTTAAGCTTACTGCG 1260

Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTTAAGCTTACTGCG 1260

QY 1261 CGAAGCGCTTGGAAATAGGCGGCTGCTGCTATATGTTATTTTCCACCATATTG 1320

Db 1261 CGAAGCGCTTGGAAATAGGCGGCTGCTGCTATATGTTATTTTCCACCATATTG 1320

QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGTCTTCTTACGAGCATTCCT 1380

Db 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGTCTTCTTACGAGCATTCCT 1380

QY 1381 AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGGTCTGTTGATGCTGTAAGAGCA 1440

Db 1381 AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGGTCTGTTGATGCTGTAAGAGCA 1440

Db 1381 AGGGGCTTTTCCCTCTCGCCCAAGGAATGCAAGGTCTGTGTAATGTCGTGAGGAGCA 1440  
QY 1441 GTTCTCTTGAAAGCTTCTTGAAGACAAAACAACTCTGTAGCGACCTTTGACAGGAGCGG 1500  
Db 1441 GTTCTCTTGAAAGCTTCTTGAAGACAAAACAACTCTGTAGCGACCTTTGAGGAGCGG 1500  
QY 1501 AACCCCCCACTCGGACAGGTGCTCTGGGCCAAAAGCCAGTGTATAGATACACCT 1560  
Db 1501 AACCCCCCACTCGGACAGGTGCTCTGGGCCAAAAGCCAGTGTATAGATACACCT 1560  
QY 1561 GCAAGGCGGCAACAACCCAGTCCAGTTGTGAGTTGAGTAGTTGGAAGAGTCAAA 1620  
Db 1561 GCAAGGCGGCAACAACCCAGTCCAGTTGTGAGTTGAGTAGTTGGAAGAGTCAAA 1620  
QY 1621 TGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT 1680  
Db 1621 TGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT 1680  
QY 1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTTAGTCGAGGTTAAAA 1740  
Db 1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTTAGTCGAGGTTAAAA 1740  
QY 1741 AAGCTCTAGGCCCCCGAAACACGGGACGTGTTTTCTTTGAAAAACACGATAATACC 1800  
Db 1741 AAGCTCTAGGCCCCCGAAACACGGGACGTGTTTTCTTTGAAAAACACGATAATACC 1800  
QY 1801 ATGGCGCTATTACGGCTACTCCCAACAGCGAGCGTCTACTTGGCTGCATCATCACT 1860  
Db 1801 ATGGCGCTATTACGGCTACTCCCAACAGCGAGCGTCTACTTGGCTGCATCATCACT 1860  
QY 1861 AGCTCACAGCGCGGACAGGACAGGTGAGGGGAGTCCAAAGTGGTCTCCACGCA 1920  
Db 1861 AGCTCACAGCGCGGACAGGACAGGTGAGGGGAGTCCAAAGTGGTCTCCACGCA 1920  
QY 1921 ACACAATCTTTCTGCGACCTCGTCAATGGGTGTGTTGGACTGTCTATATGTTGCC 1980  
Db 1921 ACACAATCTTTCTGCGACCTCGTCAATGGGTGTGTTGGACTGTCTATATGTTGCC 1980  
QY 1981 GGCTAAAGACCTTTCGGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
Db 1981 GGCTAAAGACCTTTCGGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
QY 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGGGCGGCTTCTTGACACCATGCACCTGC 2100  
Db 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGGGCGGCTTCTTGACACCATGCACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTGGTACAGAGCATGCCGATGTCATTCGGGTGCGCGGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTGGTACAGAGCATGCCGATGTCATTCGGGTGCGCGGCGG 2160  
QY 2161 GGCAGCAGCAGGGGAGCTACTCTCCCGAGGCCGTCTCTACTTGAAGGGCTCTTCG 2220  
Db 2161 GGCAGCAGCAGGGGAGCTACTCTCCCGAGGCCGTCTCTACTTGAAGGGCTCTTCG 2220  
QY 2221 GCGGCTCACTGTCTGCCCTCGGGCAGCTGTGGGCATCTTTGGGGCTGCGGTGTC 2280  
Db 2221 GCGGCTCACTGTCTGCCCTCGGGCAGCTGTGGGCATCTTTGGGGCTGCGGTGTC 2280  
QY 2281 ACCCGAGGGGTGCGAAGCGGTGGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCCGAGGGGTGCGAAGCGGTGGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGTCTTACGGCAACTCGTCCCTCGGCCGTACCGAGACATTTCCAGGTG 2400  
Db 2341 CGGTCCCGGTCTTACGGCAACTCGTCCCTCGGCCGTACCGAGACATTTCCAGGTG 2400  
QY 2401 GCCCATCTACGCCCTACTGTAGCGGCAAGACACTAAGTGGCGGTGGTAGCA 2460  
Db 2401 GCCCATCTACGCCCTACTGTAGCGGCAAGACACTAAGTGGCGGTGGTAGCA 2460  
QY 2461 GCCCAAGGGGTATAGGTGCTTGTCTGAAACCCGTGCGGCCCACTAGTGTTCGGG 2520  
Db 2461 GCCCAAGGGGTATAGGTGCTTGTCTGAAACCCGTGCGGCCCACTAGTGTTCGGG 2520

QY 2521 GCGTATATGCTTAAGGCACATGGTATCGACCCCTAAACATCAGAACCGGGTAAAGCCATC 2580  
Db 2521 GCGTATATGCTTAAGGCACATGGTATCGACCCCTAAACATCAGAACCGGGTAAAGCCATC 2580  
QY 2581 ACCACGGGTGCCCCATCAGCTACTCCACTATGGCAAGTTTCTTGCCGACGGTGGTTCG 2640  
Db 2581 ACCACGGGTGCCCCATCAGCTACTCCACTATGGCAAGTTTCTTGCCGACGGTGGTTCG 2640  
QY 2641 TCTGGGGGCGCTTATGACATCATATATGTGATGTAGTGCCACTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGGCGCTTATGACATCATATATGTGATGTAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTCTGGGATCGGACAGTCTGGACCAAGCGGACAGCGCTGGAGCGGACTCGTCTG 2760  
Db 2701 ATCTCTGGGATCGGACAGTCTGGACCAAGCGGACAGCGCTGGAGCGGACTCGTCTG 2760  
QY 2761 CTGCCACCGCTACGCTCGGGATCGGTCAACCGTGCCACATCAAAACATCGAGGAGGTG 2820  
Db 2761 CTGCCACCGCTACGCTCGGGATCGGTCAACCGTGCCACATCAAAACATCGAGGAGGTG 2820  
QY 2821 GCTCTCTCAGCACTCGAGAAATCCCTTTATGGCAAGCCATCCCAATCGAGACCATC 2880  
Db 2821 GCTCTCTCAGCACTCGAGAAATCCCTTTATGGCAAGCCATCCCAATCGAGACCATC 2880  
QY 2881 AAGGGGGGAGGACACTCTTCTGCCATTCCAAAGAAATGTGATGAGCTCGCGCG 2940  
Db 2881 AAGGGGGGAGGACACTCTTCTGCCATTCCAAAGAAATGTGATGAGCTCGCGCG 2940  
QY 2941 AAGCTGTGCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATGATCGCTC 3000  
Db 2941 AAGCTGTGCGGCTCGGACTCAATGTGTAGCATATTAACGGGGCTTGTATGATCGCTC 3000  
QY 3001 ATACCAACTAGCGGACAGCTCATTTGTCTGAGCAACCGAGCTCTAATAGCGGCTTTACC 3060  
Db 3001 ATACCAACTAGCGGACAGCTCATTTGTCTGAGCAACCGAGCTCTAATAGCGGCTTTACC 3060  
QY 3061 GCGGATTTGACTCAGTGTGATCGAATACATGATGTGACCCAGACAGTTCGACTTCAGC 3120  
Db 3061 GCGGATTTGACTCAGTGTGATCGAATACATGATGTGACCCAGACAGTTCGACTTCAGC 3120  
QY 3121 CTGGACCCGACCTTACCACTTTAGAGACAGCAGCTGTCACAAAGCGGCTGTCACGCTCG 3180  
Db 3121 CTGGACCCGACCTTACCACTTTAGAGACAGCAGCTGTCACAAAGCGGCTGTCACGCTCG 3180  
QY 3181 CAGCGCGAGGACAGGACTGGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
Db 3181 CAGCGCGAGGACAGGACTGGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGGCCCTCGGGCATGTTCCGATTCCTCGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
Db 3241 GAACGGCCCTCGGGCATGTTCCGATTCCTCGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
QY 3301 GCTTGTGTACGAGCTCACGCCCGCGGAGACTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTGTACGAGCTCACGCCCGCGGAGACTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGGTTCGCCCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTC 3420  
Db 3361 CCAGGGTTCGCCCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTC 3420  
QY 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTAAGCAGCAGGAGACAACTTCCCTTAC 3480  
Db 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTAAGCAGCAGGAGACAACTTCCCTTAC 3480  
QY 3481 CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
Db 3481 CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
QY 3541 CAAATGTGAAGTGTCTCATACGGCTAAAGCCCTACGCTGACGCGGCCAACGCGCCCTGCTG 3600  
Db 3541 CAAATGTGAAGTGTCTCATACGGCTAAAGCCCTACGCTGACGCGGCCAACGCGCCCTGCTG 3600

QY 3601 TATAGGCTGGAGCGGTTCAAAACGAGGTTACTACACACACCCCAATACCAATACATC 3660  
DB |||||  
QY 3601 TATAGGCTGGAGCGGTTCAAAACGAGGTTACTACACACACCCCAATACCAATACATC 3660  
DB |||||  
QY 3661 ATGGCANTGATGTGGCTGACCTGGAGGTGCTACAGAGCACCTGGTGTGGTGGCGGA 3720  
DB |||||  
QY 3661 ATGGCANTGATGTGGCTGACCTGGAGGTGCTACAGAGCACCTGGTGTGGTGGCGGA 3720  
DB |||||  
QY 3721 GTCTAGACGCTGTGGCGGCTATGCTGCTGACAAAGGAGCGTGTGATTTGGGCGAG 3780  
DB |||||  
QY 3721 GTCTAGACGCTGTGGCGGCTATGCTGCTGACAAAGGAGCGTGTGATTTGGGCGAG 3780  
DB |||||  
QY 3781 ATCATCTTCCGGAAGCGGCGCATCATTTCCCGACAGAGGAAGTCTTTACCGGAGTTTC 3840  
DB |||||  
QY 3781 ATCATCTTCCGGAAGCGGCGCATCATTTCCCGACAGAGGAAGTCTTTACCGGAGTTTC 3840  
DB |||||  
QY 3841 GATGAGATGGAGAGTGGGCTTACACCTCCCTTACATCGAAACAGGGAATGACGCTCGCC 3900  
DB |||||  
QY 3841 GATGAGATGGAGAGTGGGCTTACACCTCCCTTACATCGAAACAGGGAATGACGCTCGCC 3900  
DB |||||  
QY 3901 GAACAATTCAAACAGAGGCAATCGGTTGCTGCAAAAGCCACCAAGCAAGCGAGGCT 3960  
DB |||||  
QY 3901 GAACAATTCAAACAGAGGCAATCGGTTGCTGCAAAAGCCACCAAGCAAGCGAGGCT 3960  
DB |||||  
QY 3961 GCTGCTCCGCTGGTGGAAATCCAAAGTGGCGGACCTCGAAGCCCTTCTGGCGGAAGCATATG 4020  
DB |||||  
QY 3961 GCTGCTCCGCTGGTGGAAATCCAAAGTGGCGGACCTCGAAGCCCTTCTGGCGGAAGCATATG 4020  
DB |||||  
QY 4021 TGGAAATTCATCAGCGGGATCAATATTTAGCAGGCTTGCTCACTGTGCTGGCAACCCC 4080  
DB |||||  
QY 4021 TGGAAATTCATCAGCGGGATCAATATTTAGCAGGCTTGCTCACTGTGCTGGCAACCCC 4080  
DB |||||  
QY 4081 GGGATAGCATCATGATGSCATTCACAGCTCTATCACAGCCGCTCACACCCCAACAT 4140  
DB |||||  
QY 4081 GGGATAGCATCATGATGSCATTCACAGCTCTATCACAGCCGCTCACACCCCAACAT 4140  
DB |||||  
QY 4141 ACCCTCCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCAACTTGTCTCCCGAGCGCT 4200  
DB |||||  
QY 4141 ACCCTCCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCAACTTGTCTCCCGAGCGCT 4200  
DB |||||  
QY 4201 GCTTCTGCTTTGTTAGGCGCGGATCGCTGAGCGGCTGTGGCAGCATAGGCTTTGGG 4260  
DB |||||  
QY 4201 GCTTCTGCTTTGTTAGGCGCGGATCGCTGAGCGGCTGTGGCAGCATAGGCTTTGGG 4260  
DB |||||  
QY 4261 AAGTGCTTTGCGATATTTGSCAGGTTATGGAGCAGGGGTGGCAGGCGGCTCGTGCC 4320  
DB |||||  
QY 4261 AAGTGCTTTGCGATATTTGSCAGGTTATGGAGCAGGGGTGGCAGGCGGCTCGTGCC 4320  
DB |||||  
QY 4321 TTTAAGGTCTAGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 4380  
DB |||||  
QY 4321 TTTAAGGTCTAGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 4380  
DB |||||  
QY 4381 ATCTCTCCCTGGGCGCCCTAGTCTGCGGGTCTGTGGGAGCGATATCTGGCTCGGCAC 4440  
DB |||||  
QY 4381 ATCTCTCCCTGGGCGCCCTAGTCTGCGGGTCTGTGGGAGCGATATCTGGCTCGGCAC 4440  
DB |||||  
QY 4441 GTGGGCCCAAGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
DB |||||  
QY 4441 GTGGGCCCAAGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCGTTGCTTCGCGG 4500  
DB |||||  
QY 4501 GGTAACACGCTCTCCCCACGACATATGTGCTGAGAGCGAGCTGACAGCAGTGTCACT 4560  
DB |||||  
QY 4501 GGTAACACGCTCTCCCCACGACATATGTGCTGAGAGCGAGCTGACAGCAGTGTCACT 4560  
DB |||||  
QY 4561 CAGATCTCTTAGTCTTTACCATCACTCAGCTGTGAAAGGCTTCAACAGTGGATCAAC 4620  
DB |||||  
QY 4561 CAGATCTCTTAGTCTTTACCATCACTCAGCTGTGAAAGGCTTCAACAGTGGATCAAC 4620  
DB |||||  
QY 4621 GAGGACTGTCTCACGCCATGTCTCGGCTCGTGGCTAAGAGATGTTTGGATTTGATATGC 4680  
DB |||||  
QY 4621 GAGGACTGTCTCACGCCATGTCTCGGCTCGTGGCTAAGAGATGTTTGGATTTGATATGC 4680  
DB |||||  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCTCGCGGATGCGCGGA 4740

DB |||||  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCGCGGATGCGCGGA 4740  
DB |||||  
QY 4741 GTCCCTTCTTCTCATGTCAACGTGGGTAAAGGGAGTCTGGCGGGCGACGGCATCATG 4800  
DB |||||  
QY 4741 GTCCCTTCTTCTCATGTCAACGTGGGTAAAGGGAGTCTGGCGGGCGACGGCATCATG 4800  
DB |||||  
QY 4801 CAAACCACTGCGCATGTGGAGCAGATACCGGACATGTGAATAACCGTTCCATGAGG 4860  
DB |||||  
QY 4801 CAAACCACTGCGCATGTGGAGCAGATACCGGACATGTGAATAACCGTTCCATGAGG 4860  
DB |||||  
QY 4861 ATCTGCGGGCTTAGCACCTGTAGTAACAGTGGCATGGAACATTTCCCAATTAACGCGTAC 4920  
DB |||||  
QY 4861 ATCTGCGGGCTTAGCACCTGTAGTAACAGTGGCATGGAACATTTCCCAATTAACGCGTAC 4920  
DB |||||  
QY 4921 ACCACGGGCTTGTCAACGCTTCCCGGCGCAAAATTAATCTAGGCGCTGTGGCGGGTG 4980  
DB |||||  
QY 4921 ACCACGGGCTTGTCAACGCTTCCCGGCGCAAAATTAATCTAGGCGCTGTGGCGGGTG 4980  
DB |||||  
QY 4981 GCTGCTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTCCACTACGTGACGGCATG 5040  
DB |||||  
QY 4981 GCTGCTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTCCACTACGTGACGGCATG 5040  
DB |||||  
QY 5041 ACCACTGAACGTAAGTGCCTGTAGGTTCCGGGCCCCCGAAATTTCTCACAGAAGTG 5100  
DB |||||  
QY 5041 ACCACTGAACGTAAGTGCCTGTAGGTTCCGGGCCCCCGAAATTTCTCACAGAAGTG 5100  
DB |||||  
QY 5101 GATGGGTGCGGTTGCAACGATACGCTCAGCGTCAAAACCCCTCTACGGGAGGAGTTC 5160  
DB |||||  
QY 5101 GATGGGTGCGGTTGCAACGATACGCTCAGCGTCAAAACCCCTCTACGGGAGGAGTTC 5160  
DB |||||  
QY 5161 ACATTCCTGTCGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCCGAA 5220  
DB |||||  
QY 5161 ACATTCCTGTCGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCCGAA 5220  
DB |||||  
QY 5221 CCGGACGTAGCAGTCTCACTTCCATGTCTACCGACCCCTCCACATACGGCGGAGAG 5280  
DB |||||  
QY 5221 CCGGACGTAGCAGTCTCACTTCCATGTCTACCGACCCCTCCACATACGGCGGAGAG 5280  
DB |||||  
QY 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTCTTGGCGAGCTCATCAGTAGCCAG 5340  
DB |||||  
QY 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTCTTGGCGAGCTCATCAGTAGCCAG 5340  
DB |||||  
QY 5341 CTGCTGCGCTTCTTGAAGCAACATGCACTACCCGTCTGCTCCCGGACGCTGAC 5400  
DB |||||  
QY 5341 CTGCTGCGCTTCTTGAAGCAACATGCACTACCCGTCTGCTCCCGGACGCTGAC 5400  
DB |||||  
QY 5401 CTCATCGAGGCAACCTCTCTGTCGGCGAGAGATGGCGGGAAACATCACCCGCTGGAG 5460  
DB |||||  
QY 5401 CTCATCGAGGCAACCTCTCTGTCGGCGAGAGATGGCGGGAAACATCACCCGCTGGAG 5460  
DB |||||  
QY 5461 TCAGAAATATAGGTAGTAAATTTTGAATCTTTTCGAGCCGCTCCAGCGGAGGAGATGAG 5520  
DB |||||  
QY 5461 TCAGAAATATAGGTAGTAAATTTTGAATCTTTTCGAGCCGCTCCAGCGGAGGAGATGAG 5520  
DB |||||  
QY 5521 AGGGAAGTATCCGTTCCGCGGAGATCTGCGGAGGTCCAGAAATTCCTCTGAGCGATG 5580  
DB |||||  
QY 5521 AGGGAAGTATCCGTTCCGCGGAGATCTGCGGAGGTCCAGAAATTCCTCTGAGCGATG 5580  
DB |||||  
QY 5581 CCCATATGGGCAACGCGGATTAACAAACCTCTCCACTGTTAGAGTCTGGAAGGACCCGGAC 5640  
DB |||||  
QY 5581 CCCATATGGGCAACGCGGATTAACAAACCTCTCCACTGTTAGAGTCTGGAAGGACCCGGAC 5640  
DB |||||  
QY 5641 TAGCTCCCTCAGTGGTACAGGGTGTCCATTCGCGCTGCAAGGCCCCCTCCCGATACA 5700  
DB |||||  
QY 5641 TAGCTCCCTCAGTGGTACAGGGTGTCCATTCGCGCTGCAAGGCCCCCTCCCGATACA 5700  
DB |||||  
QY 5701 CCTCCAGGAGAGAGGACGCTTGTCTGTGAGAACTCTACGCTGTCTTCTGCTTTGGG 5760  
DB |||||  
QY 5701 CCTCCAGGAGAGAGGACGCTTGTCTGTGAGAACTCTACGCTGTCTTCTGCTTTGGG 5760  
DB |||||  
QY 5761 GAGTCTGCGCAAAAGACCTTCGGGAGCTCCGAAATGCTCGGCGCTCGAACGGGACCGGA 5820

Db 5761 GAGTCGCCACAAGACCTTGGGAGCTCCGAACTGTCGGCCGTCGACAGCGGCACGGCA 5820  
Qy 5821 ACGGCTCTCTGACACAGCCCTCCGACGACGGCGACCGCGGATCCGACGTTGAGTCGTAC 5880  
Db 5821 ACGGCTCTCTGACACAGCCCTCCGACGACGGCGACCGCGGATCCGACGTTGAGTCGTAC 5880  
Qy 5881 TCTTCATGCCCCCTTGAAGGGGAGCGGGGATCCGATCTCAGACGCGGTTCTGG 5940  
Db 5881 TCTTCATGCCCCCTTGAAGGGGAGCGGGGATCCGATCTCAGACGCGGTTCTGG 5940  
Qy 5941 TCTACCGTAAGCAGAGAGCTAGTGAGGACGTCGTCTGCTGCTGATGTCCTACACATGG 6000  
Db 5941 TCTACCGTAAGCAGAGAGCTAGTGAGGACGTCGTCTGCTGCTGATGTCCTACACATGG 6000  
Qy 6001 ACAGGCGCCTGATCAGCCATGCGTGGGAGGAAACCAAGTCGCCATCAATCAACTG 6060  
Db 6001 ACAGGCGCCTGATCAGCCATGCGTGGGAGGAAACCAAGTCGCCATCAATCAACTG 6060  
Qy 6061 AGCAACTCTTGTCCGTCACCAACACTTGGTCTATGCTACACATCTCGCAGCAAGC 6120  
Db 6061 AGCAACTCTTGTCCGTCACCAACACTTGGTCTATGCTACACATCTCGCAGCAAGC 6120  
Qy 6121 CTGCGCAGAGAAGGTCACTTTGACAGACTCGAGTCTGAGCAGCACTACCGGAC 6180  
Db 6121 CTGCGCAGAGAAGGTCACTTTGACAGACTCGAGTCTGAGCAGCACTACCGGAC 6180  
Qy 6181 GTGCTCAAGAGATGAAGGCAAGGCGTCCACAGTTAAGGCTAAACTTCTATCCGTGGAG 6240  
Db 6181 GTGCTCAAGAGATGAAGGCAAGGCGTCCACAGTTAAGGCTAAACTTCTATCCGTGGAG 6240  
Qy 6241 GAAGCTGTAAAGCTGAGCGCCCACTTCGAGAGGCGTTAAACACATCGCTCCGTTGGAAGCACTT 6300  
Db 6241 GAAGCTGTAAAGCTGAGCGCCCACTTCGAGAGGCGTTAAACACATCGCTCCGTTGGAAGCACTT 6300  
Qy 6301 GAGCTCGGAACCTATCCAGCAAGGCGTTAAACACATCGCTCCGTTGGAAGCACTT 6360  
Db 6301 GAGCTCGGNACTATCCAGCAAGGCGTTAAACACATCGCTCCGTTGGAAGCACTT 6360  
Qy 6361 CTGGAAGACTGAGACACCAATTTGACACCAACCATCATGGCAAAAATAGGTTTCTGC 6420  
Db 6361 CTGGAAGACTGAGACACCAATTTGACACCAACCATCATGGCAAAAATAGGTTTCTGC 6420  
Qy 6421 GTCCAAACAGAGAGGGGCGCAGCAGCTCGCTTATCGATATCCAGATTTGGGG 6480  
Db 6421 GTCCAAACAGAGAGGGGCGCAGCAGCTCGCTTATCGATATCCAGATTTGGGG 6480  
Qy 6481 GTTCGTGTGCGAGAAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCGGTG 6540  
Db 6481 GTTCGTGTGCGAGAAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCGGTG 6540  
Qy 6541 ATGGCTCTTCAACGATTTCCAAATCTCTCTGGAAGCGGTTCGAGTTCTGTGTAAT 6600  
Db 6541 ATGGCTCTTCAACGATTTCCAAATCTCTCTGGAAGCGGTTCGAGTTCTGTGTAAT 6600  
Qy 6601 GCTCGAAGCAGAGAAATGCGCTATGGCTTCGATATGACACCGCTGTTTGACTCA 6660  
Db 6601 GCTCGAAGCAGAGAAATGCGCTATGGCTTCGATATGACACCGCTGTTTGACTCA 6660  
Qy 6661 ACAGTCTACTGAGATGACATCCGTTGAGGAGTCAATCTACCAATGTTGTGACTGGCC 6720  
Db 6661 ACAGTCTACTGAGATGACATCCGTTGAGGAGTCAATCTACCAATGTTGTGACTGGCC 6720  
Qy 6721 CCGGAGCCAGACAGGCCATAAGGTCGCTCACAGAGCGGTTTACATCGGGGCGCCCTG 6780  
Db 6721 CCGGAGCCAGACAGGCCATAAGGTCGCTCACAGAGCGGTTTACATCGGGGCGCCCTG 6780  
Qy 6781 ACTAAATCTAAAGGCAAGACTCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
Db 6781 ACTAAATCTAAAGGCAAGACTCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
Qy 6841 ACCAGCTGCGGTAAATACCTTCACATGTTACTTGAAGCCGCTGCGGCTCTCGAGCTGCG 6900  
Db 6841 ACCAGCTGCGGTAAATACCTTCACATGTTACTTGAAGCCGCTGCGGCTCTCGAGCTGCG 6900

Qy 6901 AAGTCCAGGACTGCAAGATGCTCGTATCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960  
Db 6901 AAGTCCAGGACTGCAAGATGCTCGTATCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960  
Qy 6961 GCGGGACCCAAAGAGAGAGCGAGCTTACGGGCTTCAAGGAGCTATGACTAGATAC 7020  
Db 6961 GCGGGACCCAAAGAGAGAGCGAGCTTACGGGCTTCAAGGAGCTATGACTAGATAC 7020  
Qy 7021 TCTGCCCCCTCTGGGACCCGCCAAACAGAAATACGATTTGGAGTTGATAACATCATGC 7080  
Db 7021 TCTGCCCCCTCTGGGACCCGCCAAACAGAAATACGATTTGGAGTTGATAACATCATGC 7080  
Qy 7081 TCTTCCAATGTGTGCTCGGACAGTGCATCTGCGCAAAAGGTTGACTATCTCAACCGT 7140  
Db 7081 TCTTCCAATGTGTGCTCGGACAGTGCATCTGCGCAAAAGGTTGACTATCTCAACCGT 7140  
Qy 7141 GACCCACACACCCCTTGGCGGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
Db 7141 GACCCACACACCCCTTGGCGGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
Qy 7201 TCTTGGCTAGGCAACATCATGATGCGCCACCTTGTGGCAAGGATGATCTGTATG 7260  
Db 7201 TCTTGGCTAGGCAACATCATGATGCGCCACCTTGTGGCAAGGATGATCTGTATG 7260  
Qy 7261 ACTCATTTCTTCCATCTTCTAGCTCAGGACAACTTTGAAAAGCCCTAGATTGTCAG 7320  
Db 7261 ACTCATTTCTTCCATCTTCTAGCTCAGGAACTTGTGAAAAGCCCTAGATTGTCAG 7320  
Qy 7321 ATCTAGCGGCGCTGTCTTCTTCTAGGACCACTTGTGACCTCAGATCATTTCAACGACTC 7380  
Db 7321 ATCTAGCGGCGCTGTCTTCTTCTAGGACCACTTGTGACCTCAGATCATTTCAACGACTC 7380  
Qy 7381 CATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
Db 7381 CATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
Qy 7441 TATGCTCAGGAACTTGGGTTACCGCTTGTGAGTCTGAGACATCGGCGCAAGT 7500  
Db 7441 TATGCTCAGGAACTTGGGTTACCGCTTGTGAGTCTGAGACATCGGCGCAAGT 7500  
Qy 7501 GTCCGCTAGGCTACTGTCCAGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Db 7501 GTCCGCTAGGCTACTGTCCAGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Qy 7561 AACTGGGCAATAGGACCAAGCTCAAACTCACTCCAAATCCCGCTCGTCCAGTTGGAT 7620  
Db 7561 AACTGGGCAATAGGACCAAGCTCAAACTCACTCCAAATCCCGCTCGTCCAGTTGGAT 7620  
Qy 7621 TATCCAGCTGTGTTCTGCTGTACAGGGGGAGACATATACAGCCCTGTCTCGT 7680  
Db 7621 TATCCAGCTGTGTTCTGCTGTACAGGGGGAGACATATACAGCCCTGTCTCGT 7680  
Qy 7681 GCGGACCCCGCTGTGTTCTGCTACTCTACTTCTAGGGGTAGGCACTAT 7740  
Db 7681 GCGGACCCCGCTGTGTTCTGCTACTCTACTTCTAGGGGTAGGCACTAT 7740  
Qy 7741 CTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Db 7741 CTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Qy 7801 TTTCCCTTT 7860  
Db 7801 TTTCCCTTT 7860  
Qy 7861 TTTTCTCTTT 7920  
Db 7861 TTTTCTCTTT 7920  
Qy 7921 TAGCTGTGAAGGTCCGTGAGCGCTTGTGCTGAGAGAGTGTGATCTGGCCTCTCTGC 7980  
Db 7921 TAGCTGTGAAGGTCCGTGAGCGCTTGTGCTGAGAGAGTGTGATCTGGCCTCTCTGC 7980

```
QY 7981 AGATCAAGT 7989
|||||
Db 7981 AGATCAAGT 7989

RESULT 4
AX739972
LOCUS AX739972 10690 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 3 from Patent WO20059321.
ACCESSION AX739972
VERSION AX739972.1 GI:30519246
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS de Francesco, R., Migliaccio, G. and Paonessa, G.
TITLE Hepatitis C virus replicons and replicon enhanced cells
JOURNAL Patent: WO 02059321-A 3 01-AUG-2002;
INSTITUTO DI RICERCHE DI BIOLOGIA MOLECOLARE P. ANGELETTI S.P.A.
(IT)

FEATURES
Location/Qualifiers
source 1..10690
/mol_type="unidentified"
/db_xref="taxon:32644"
/note="pHCVNeo.17 coding"

ORIGIN
Query Match 100.0%; Score 7987.4; DB 6; Length 10690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATTCGGGGGACACCTCCACCATAGATCACTCCCTGTGAGGAATCTGTG 60
Db 1 GCCAGCCCCGATTCGGGGGACACCTCCACCATAGATCACTCCCTGTGAGGAATCTGTG 60

QY 61 TCTTCACGACGAGTGTAGCCATGCGTGTAGTATGAGTGTGCTGAGCCTCCAGGAC 120
Db 61 TCTTCACGACGAGTGTAGCCATGCGTGTAGTATGAGTGTGCTGAGCCTCCAGGAC 120

QY 121 CCCCGCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180
Db 121 CCCCGCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180

QY 181 GACGACCGGTTCCTTTCTTTGATCAACCGCTCAATGCTGAGATTTGGGGTCCCGCC 240
Db 181 GACGACCGGTTCCTTTCTTTGATCAACCGCTCAATGCTGAGATTTGGGGTCCCGCC 240

QY 241 GCGAGACTGTAGCGGAGTGTGTTGGGTGCGAAAGGCTTGTGGTACTGCTGATAGG 300
Db 241 GCGAGACTGTAGCGGAGTGTGTTGGGTGCGAAAGGCTTGTGGTACTGCTGATAGG 300

QY 301 GTGCTTGCGAGTCCCGGGAGGTCTCGTAGACGCTGACCATGAGCAGCAATCTTAAC 360
Db 301 GTGCTTGCGAGTCCCGGGAGGTCTCGTAGACGCTGACCATGAGCAGCAATCTTAAC 360

QY 361 CTCAAGAAAAACAAAGGCGCGCATGATTGACAAAGATGGATTGCAACGACGTTCTC 420
Db 361 CTCAAGAAAAACAAAGGCGCGCATGATTGACAAAGATGGATTGCAACGACGTTCTC 420

QY 421 CGGCGCTTGGGTGGAGAGGCTATTTCGCTATGATGGGACAAACAGCAATCGGCTGCT 480
Db 421 CGGCGCTTGGGTGGAGAGGCTATTTCGCTATGATGGGACAAACAGCAATCGGCTGCT 480

QY 481 CTGATGCGCGGTGTTCCGGCTGTGACGCGAGGGGCGCCGGTCTTTTGTCAAGACCG 540
Db 481 CTGATGCGCGGTGTTCCGGCTGTGACGCGAGGGGCGCCGGTCTTTTGTCAAGACCG 540

QY 541 ACCTGTCCGGTCCCTGAAATGAATGCAATGTCAGGACGAGGCGCGGTATCGTGGCTGGCCA 600
Db 541 ACCTGTCCGGTCCCTGAAATGAATGCAATGTCAGGACGAGGCGCGGTATCGTGGCTGGCCA 600

QY 601 CGACGGGGCTTCCTTGGCGCAGCTGTCTGCGAGTGTCTCACTGAACGGGAAGGACTGGC 660
Db 601 CGACGGGGCTTCCTTGGCGCAGCTGTCTGCGAGTGTCTCACTGAACGGGAAGGACTGGC 660

QY 661 TGCTATTGGCGAAGTGTCCGGGCGAGGATCTCTCTGTCATCTCACCTTGTCTCTGCGAGA 720
Db 661 TGCTATTGGCGAAGTGTCCGGGCGAGGATCTCTCTGTCATCTCACCTTGTCTCTGCGAGA 720

QY 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCATACGCTTGTATCGGCTACCTGGC 780
Db 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCATACGCTTGTATCGGCTACCTGGC 780

QY 781 CATTCGACCAACCAAGCGAAACATCGCATCGACGAGCAGTACTCGATGGAAGCGGTC 840
Db 781 CATTCGACCAACCAAGCGAAACATCGCATCGACGAGCAGTACTCGATGGAAGCGGTC 840

QY 841 TTGTCGATCAGGATGTGACGAGAGCATCAGGGGCTCGCGGCGAGCCGAATCTTTCG 900
Db 841 TTGTCGATCAGGATGTGACGAGAGCATCAGGGGCTCGCGGCGAGCCGAATCTTTCG 900

QY 901 CCAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGTGTGACCATGCGCATGCT 960
Db 901 CCAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGTGTGACCATGCGCATGCT 960

QY 961 GCTTCCCGAATATCATGTGGAATATGCGCTTTTCTGGATTCTATCGACTGTGGCGCGC 1020
Db 961 GCTTCCCGAATATCATGTGGAATATGCGCTTTTCTGGATTCTATCGACTGTGGCGCGC 1020

QY 1021 TGGGTGTGCGGACCGCTATCAGGACATGAGGCTTGGCTACCGGTGATATTGCTGAAGAGC 1080
Db 1021 TGGGTGTGCGGACCGCTATCAGGACATGAGGCTTGGCTACCGGTGATATTGCTGAAGAGC 1080

QY 1081 TTGGCGGCGAATGGGCTGACCGCTTCTGCTGCTTTTACGGTATCCCGCTGCCGATTCGC 1140
Db 1081 TTGGCGGCGAATGGGCTGACCGCTTCTGCTGCTTTTACGGTATCCCGCTGCCGATTCGC 1140

QY 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTCTTCTGAGTCTTCTGAGTCT 1200
Db 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTCTTCTGAGTCTTCTGAGTCT 1200

QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAAGTACTGGC 1260
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAAGTACTGGC 1260

QY 1261 CGAAGCGCTTGGATAGCGCGGTGTGCTGCTTCTATATGTTATTTTCCACCATTTG 1320
Db 1261 CGAAGCGCTTGGATAGCGCGGTGTGCTGCTTCTATATGTTATTTTCCACCATTTG 1320

QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTCTTACGAGCATTCCT 1380
Db 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTCTTACGAGCATTCCT 1380

QY 1381 AGGGGTCTTCCCTCTGCGCAAGGAATGCAAGGTCTGTTGAATGTCTGTAAGGAAGCA 1440
Db 1381 AGGGGTCTTCCCTCTGCGCAAGGAATGCAAGGTCTGTTGAATGTCTGTAAGGAAGCA 1440

QY 1441 GTTCTCTGGAAGCTTCTTGAAGCAACACGCTCTGAGGACCTTTCGAGGACGG 1500
Db 1441 GTTCTCTGGAAGCTTCTTGAAGCAACACGCTCTGAGGACCTTTCGAGGACGG 1500

QY 1501 AACCCCCCCTGCGGACAGGTGCTCTGCGGCGCAAGGACCGCTGTATAGATACACCT 1560
Db 1501 AACCCCCCCTGCGGACAGGTGCTCTGCGGCGCAAGGACCGCTGTATAGATACACCT 1560

QY 1561 GCAAAGGCGGACAAACCCAGTGTGCTGAGTGTGGAATGTGTTGGAAGAGTCAAA 1620
Db 1561 GCAAAGGCGGACAAACCCAGTGTGCTGAGTGTGGAATGTGTTGGAAGAGTCAAA 1620

QY 1621 TGGCTCTCTCAAGCGTATTCAACAAGGGCTGAGGATGCCAGAGGTACCCCATTTG 1680
Db 1621 TGGCTCTCTCAAGCGTATTCAACAAGGGCTGAGGATGCCAGAGGTACCCCATTTG 1680

QY 1681 ATGGGATCTGATCTGGGGCCTCGGTGCACATGCTTTTACATGTTTGTAGTTCAGGTTAAA 1740
```



1681 ATGGGATCTGATCTGGGGCTCGGTGCACATCTTTACATGTTGTAGTCGAGGTTAAAA 1740  
1741 AACGCTCAGGCCCGCCGAAACACAGGGGAGCTGTTTCTTTGAAAAACACAGTAATACC 1800  
1741 AACGCTCAGGCCCGCCGAAACACAGGGGAGCTGTTTCTTTGAAAAACACAGTAATACC 1800  
1801 ATGGGCGCTATTACGGCCTACTCCCAACAGACGCGAGGCGCTACTTGGCTGCATCATCACT 1860  
1801 ATGGGCGCTATTACGGCCTACTCCCAACAGACGCGAGGCGCTACTTGGCTGCATCATCACT 1860  
1861 AGCCTCAGAGCGCGGACAGGAAACAGGTCGAGGGGAGGTCAGAGTGCCTCCACCGCA 1920  
1861 AGCCTCAGAGCGCGGACAGGAAACAGGTCGAGGGGAGGTCAGAGTGCCTCCACCGCA 1920  
1921 ACACAACTCTTCTCGGCGACCTCGCTCAATGGCGTGTGTGGACTGTCTATCATGGTGCC 1980  
1921 ACACAACTCTTCTCGGCGACCTCGCTCAATGGCGTGTGTGGACTGTCTATCATGGTGCC 1980  
1981 GGCCTAAAGACGCTTCCCGGCCCAAGGCGCAATCACCCAAATGTACCAATGTGGAC 2040  
1981 GGCCTAAAGACGCTTCCCGGCCCAAGGCGCAATCACCCAAATGTACCAATGTGGAC 2040  
2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGCGCTTCTTGACACATGCACCTGC 2100  
2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGCGCTTCTTGACACATGCACCTGC 2100  
2101 GGCAGCTCGGACCTTTACTTGGTCAGAGGCATGCCGATGTCAATCCGGTGGCGCGCGG 2160  
2101 GGCAGCTCGGACCTTTACTTGGTCAGAGGCATGCCGATGTCAATCCGGTGGCGCGCGG 2160  
2161 GGCAGCAGAGGGGAGCCTACTCTCCCGCAGCGCGCTCTCTTAAGGGCTCTTCG 2220  
2161 GGCAGCAGAGGGGAGCCTACTCTCCCGCAGCGCGCTCTCTTAAGGGCTCTTCG 2220  
2221 GCGGTCCTACTGCTCGCCCTCGGGGACGCTGTGGGATCTTTCGGGTCGCGGTGC 2280  
2221 GCGGTCCTACTGCTCGCCCTCGGGGACGCTGTGGGATCTTTCGGGTCGCGGTGC 2280  
2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGCTCGAGTCTATGGAACCACTATG 2340  
2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGCTCGAGTCTATGGAACCACTATG 2340  
2341 CGGTCCCGCTCTACAGCAAACTCGTCCCTCGCGCGCTACCGCAGACATTCAGGTG 2400  
2341 CGGTCCCGCTCTACAGCAAACTCGTCCCTCGCGCGCTACCGCAGACATTCAGGTG 2400  
2401 GCCCATCTACGCGCCTACTGCTAGCGGACAGCACTAAGGTGCGGCTGGTATGCA 2460  
2401 GCCCATCTACGCGCCTACTGCTAGCGGACAGCACTAAGGTGCGGCTGGTATGCA 2460  
2461 GCCAAGGGTATAAGTGTGTCTCTGAACCCGCTCGCGCCACCTAGGTTCGGG 2520  
2461 GCCAAGGGTATAAGTGTGTCTCTGAACCCGCTCGCGCCACCTAGGTTCGGG 2520  
2521 GGGTATATGCTAAGGCAATGTTATCGACCTTAAATCAGAAACCGGGGTAAAGCAATC 2580  
2521 GGGTATATGCTAAGGCAATGTTATCGACCTTAAATCAGAAACCGGGGTAAAGCAATC 2580  
2581 ACCAGGGTGCCTCATCAGTACTCCACTATGCGCAAGTTTCTTCCGACGGTGTGC 2640  
2581 ACCAGGGTGCCTCATCAGTACTCCACTATGCGCAAGTTTCTTCCGACGGTGTGC 2640  
2641 TCTGGGCGCCCTATGACATATAATGTATGAGTGGCACTCAACTGACTCGACCACT 2700  
2641 TCTGGGCGCCCTATGACATATAATGTATGAGTGGCACTCAACTGACTCGACCACT 2700  
2701 ATCTTGGGCAATCGGCAAGTCTTGACCAAGCGGAGACGGCTGGAGCTCGTGTG 2760  
2701 ATCTTGGGCAATCGGCAAGTCTTGACCAAGCGGAGACGGCTGGAGCTCGTGTG 2760  
2761 CTGCGCACGCTACGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG 2820

2761 CTCGCCACGGCTACGCTCCGGGATCGGTACCGTGCCACATCCAAACATCGAGAGGTG 2820  
2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGCAAGCCATCCCATCGAGACCATC 2880  
2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGCAAGCCATCCCATCGAGACCATC 2880  
2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCGAAGAGAAATGTGATGAGCTCGCGGG 2940  
2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCGAAGAGAAATGTGATGAGCTCGCGGG 2940  
2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTCAATGATCCGTC 3000  
2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTCAATGATCCGTC 3000  
3001 ATACCAACTAGCGGAGACGTCAATGCTGTAGCAACGGACGCTCTAATGACGGGCTTACG 3060  
3001 ATACCAACTAGCGGAGACGTCAATGCTGTAGCAACGGACGCTCTAATGACGGGCTTACG 3060  
3061 GGGGATTTTCGACTCAGTGCATGACCTGATGATGTGTCACCCAGACAGCTCCGACTTCAGC 3120  
3061 GGGGATTTTCGACTCAGTGCATGATGATGTGTCACCCAGACAGCTCCGACTTCAGC 3120  
3121 CTGGACCCGACCTTCAACATTCAGACGACGACCGCTGCCAAGACGCGGTGTCAACCTCG 3180  
3121 CTGGACCCGACCTTCAACATTCAGACGACGACCGCTGCCAAGACGCGGTGTCAACCTCG 3180  
3181 CAGCGCGAGGAGGACGTGTTAGGGGACGAGTGGGCAATTTAAGGTTTGTGACTCCAGGA 3240  
3181 CAGCGCGAGGAGGACGTGTTAGGGGACGAGTGGGCAATTTAAGGTTTGTGACTCCAGGA 3240  
3241 GAACGGCTTCGGGACATGTTGATTTCTCGGTCTCTGCGAGTGTATGACCGGGCTCT 3300  
3241 GAACGGCTTCGGGACATGTTGATTTCTCGGTCTCTGCGAGTGTATGACCGGGCTCT 3300  
3301 GCTTGTAGCAGCTCAGCGCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
3301 GCTTGTAGCAGCTCAGCGCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
3361 CAGGGTTGCCGCTGTGCCAGACCATCTGAGATTCTGGGAGAGCGTCTTTACAGGCTC 3420  
3361 CAGGGTTGCCGCTGTGCCAGACCATCTGAGATTCTGGGAGAGCGTCTTTACAGGCTC 3420  
3421 ACCCATAGAGCCCATTTCTTGTCCAGACTAAGCAGCAGGAGACACTCCCTAC 3480  
3421 ACCCATAGAGCCCATTTCTTGTCCAGACTAAGCAGCAGGAGACACTCCCTAC 3480  
3481 CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
3481 CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
3541 CAAATGTGAAGTGTCTCATACGGCTTAAAGCTTACGCTGCAAGGGCAACCGCCCTGCTG 3600  
3541 CAAATGTGAAGTGTCTCATACGGCTTAAAGCTTACGCTGCAAGGGCAACCGCCCTGCTG 3600  
3601 TATAGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACCCCATTAACCAATACATC 3660  
3601 TATAGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACCCCATTAACCAATACATC 3660  
3661 ATGGCATGATGTGCGTGCATCTGAGGTCGTACAGGACCTGGGTGTGGTAGCGGA 3720  
3661 ATGGCATGATGTGCGTGCATCTGAGGTCGTACAGGACCTGGGTGTGGTAGCGGA 3720  
3721 GTCTTAGCAGCTCTGCGCGGTATTGCTGTGAACAGGAGGCGGTGTGATGTGGGACAG 3780  
3721 GTCTTAGCAGCTCTGCGCGGTATTGCTGTGAACAGGAGGCGGTGTGATGTGGGACAG 3780  
3781 ATCATCTTCTCGGAAGCGGCCATCTTCCGACAGGAGTCTTTTACCGGGAGTTC 3840  
3781 ATCATCTTCTCGGAAGCGGCCATCTTCCGACAGGAGTCTTTTACCGGGAGTTC 3840  
3841 GATGAGATGAGAGTGGCGCTCACAACCTCCCTTACATCGAACAGGGAATGACAGTTCGCC 3900  
3841 GATGAGATGAGAGTGGCGCTCACAACCTCCCTTACATCGAACAGGGAATGACAGTTCGCC 3900

QY 3901 GAACAATTCACACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
Db 3901 GAACAATTCACACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCT 3960  
QY 3961 GCTGCTCCGTTGGAATCCAAAGTGGCGGACCCCTCGAAGCCCTCTGGGCGGAAGCATATG 4020  
Db 3961 GCTGCTCCGTTGGAATCCAAAGTGGCGGACCCCTCGAAGCCCTCTGGGCGGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGGATACAATATTTAGCAGGTTGTCCACTCTGCTGGCAACCC 4080  
Db 4021 TGGAAATTCATCAGCGGGATACAATATTTAGCAGGTTGTCCACTCTGCTGGCAACCC 4080  
QY 4081 GCGATAGCATCACTGATGCAATTCACAGCTCTATCACAGCCCGCTCACACCAACAT 4140  
Db 4081 GCGATAGCATCACTGATGCAATTCACAGCTCTATCACAGCCCGCTCACACCAACAT 4140  
QY 4141 ACCCTCTCTGTTTAAACATCTGCGGGGATGGGTGGCGCCCAACTTGTCTCCACAGCGCT 4200  
Db 4141 ACCCTCTCTGTTTAAACATCTGCGGGGATGGGTGGCGCCCAACTTGTCTCCACAGCGCT 4200  
QY 4201 GCTTCTGCTTTCGTAGGCGCGGATCGCTGAGCGGCTGTGGCAGCATAGGCGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGGCGCGGATCGCTGAGCGGCTGTGGCAGCATAGGCGCTTGGG 4260  
QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGGCAGG3GTGGCAGGCGCGCTCGTGCC 4320  
Db 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGGCAGG3GTGGCAGGCGCGCTCGTGCC 4320  
QY 4321 TTTAAGGTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 4380  
Db 4321 TTTAAGGTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 4380  
QY 4381 ATCTCTCCCTGGCGGCTAGTCTCGGGGTCGTGCGGACGATACTGCGTGGCAC 4440  
Db 4381 ATCTCTCCCTGGCGGCTAGTCTCGGGGTCGTGCGGACGATACTGCGTGGCAC 4440  
QY 4441 GTGGGCCACAGGAGGGGCTGTGCAGTGGATGAACCG3GTGATAGCGTTGCTTTCGCGG 4500  
Db 4441 GTGGGCCACAGGAGGGGCTGTGCAGTGGATGAACCG3GTGATAGCGTTGCTTTCGCGG 4500  
QY 4501 GGTAAACCACTCTCCCCACGCACTATGTGCTGAGAGCGAGCTGCAAGCAGTGTCACT 4560  
Db 4501 GGTAAACCACTCTCCCCACGCACTATGTGCTGAGAGCGAGCTGCAAGCAGTGTCACT 4560  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGGAATCAAC 4620  
Db 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGGAATCAAC 4620  
QY 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATATGC 4680  
Db 4621 GAGGACTGCTCCAGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATATGC 4680  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCCTGCGCGGATTGGCGGGA 4740  
Db 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCCTGCGCGGATTGGCGGGA 4740  
QY 4741 GTCCCTCTTCTCATGTCAAGTGGGTACAAAGGAGTCTGCGGGGCGAGCGCATCATG 4800  
Db 4741 GTCCCTCTTCTCATGTCAAGTGGGTACAAAGGAGTCTGCGGGGCGAGCGCATCATG 4800  
QY 4801 CAAACCACTGCCCATGTGGAGCAGATCACCGGACATGTGAAACCGGTTCCATGAGG 4860  
Db 4801 CAAACCACTGCCCATGTGGAGCAGATCACCGGACATGTGAAACCGGTTCCATGAGG 4860  
QY 4861 ATCTGGGGGCTAGACCTGTAGTAAACGTTGGCATGGAAATTTCCCATTAACGCGTAC 4920  
Db 4861 ATCTGGGGGCTAGACCTGTAGTAAACGTTGGCATGGAAATTTCCCATTAACGCGTAC 4920  
QY 4921 ACCAGGCGCCCTGACGCGCCCTCCCGCGGCAAAATTTCTAGGGGCTGTGGCGGCTG 4980  
Db 4921 ACCAGGCGCCCTGACGCGCCCTCCCGCGGCAAAATTTCTAGGGGCTGTGGCGGCTG 4980

QY 4981 GCTGCTGAGGAGTACGTGAGGTTACGCGGTTGGGGATTTTCCACTACGTACGCGGCATG 5040  
Db 4981 GCTGCTGAGGAGTACGTGAGGTTACGCGGTTGGGGATTTTCCACTACGTACGCGGCATG 5040  
QY 5041 ACCACTGACAAAGTAAAGTGGCCGTTGTCAGGTTCCGGCCCCCGAATTTCTTACACAAGTG 5100  
Db 5041 ACCACTGACAAAGTAAAGTGGCCGTTGTCAGGTTCCGGCCCCCGAATTTCTTACACAAGTG 5100  
QY 5101 GATGGGTTGCGGTTGTCACAGGTAACGCTCCAGGTCGAAACCCCTCTTACGGGAGGAGTGC 5160  
Db 5101 GATGGGTTGCGGTTGTCACAGGTAACGCTCCAGGTCGAAACCCCTCTTACGGGAGGAGTGC 5160  
QY 5161 ACATTTCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGTCCCATCGAGCGCCGAA 5220  
Db 5161 ACATTTCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGTCCCATCGAGCGCCGAA 5220  
QY 5221 CCGGACGTAGCAGTGTCTCACTTCCATGCTACCCGACCCCTCCCATTTACGGCGGAGAGC 5280  
Db 5221 CCGGACGTAGCAGTGTCTCACTTCCATGCTACCCGACCCCTCCCATTTACGGCGGAGAGC 5280  
QY 5281 GCTAAGCGTAGGCTGGCCAGGGATCTCCGCCCTCTTGGCCAGCTCATACGTACGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGGCCAGGGATCTCCGCCCTCTTGGCCAGCTCATACGTACGCCAG 5340  
QY 5341 CTGCTCGGCTTCTTGAAGGCAACATGCACTACCCGTCATGACTCCCCCGACGCTGAC 5400  
Db 5341 CTGCTCGGCTTCTTGAAGGCAACATGCACTACCCGTCATGACTCCCCCGACGCTGAC 5400  
QY 5401 CTCATCGAGGCAACCTCTCTGTGGCGGAGGATGGCGGGAAACATCACCCCGCTGGAG 5460  
Db 5401 CTCATCGAGGCAACCTCTCTGTGGCGGAGGATGGCGGGAAACATCACCCCGCTGGAG 5460  
QY 5461 TCAGAAATAGGTAGTAAATTTTGGACTCTTTTGGAGCCGCTCCAAAGCGGAGGAGATGAG 5520  
Db 5461 TCAGAAATAGGTAGTAAATTTTGGACTCTTTTGGAGCCGCTCCAAAGCGGAGGAGATGAG 5520  
QY 5521 AGGGAAGTATCCGTTCCGGCGGAGATCTGCGGAGTCCAGGAAATTCCTCTGACGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCGGCGGAGATCTGCGGAGTCCAGGAAATTCCTCTGACGCGATG 5580  
QY 5581 CCCATATGGGCAACCGCCGATTTACAAACCTCCACTGTTTAGAGTCTCGAAGGACCCGGAC 5640  
Db 5581 CCCATATGGGCAACCGCCGATTTACAAACCTCCACTGTTTAGAGTCTCGAAGGACCCGGAC 5640  
QY 5641 TAGTCCCTCAGTGGTACACGGGTGTCATTCGCGCTCCAAAGGCGCTTCCGATACCA 5700  
Db 5641 TAGTCCCTCAGTGGTACACGGGTGTCATTCGCGCTCCAAAGGCGCTTCCGATACCA 5700  
QY 5701 CCTCCAGGAGAGAGGAGCGGTTGCTGTCAGAACTACCGTGTCTTCTGCTTGGCG 5760  
Db 5701 CCTCCAGGAGAGAGGAGCGGTTGCTGTCAGAACTACCGTGTCTTCTGCTTGGCG 5760  
QY 5761 GAGTCCGCCACAAAGACCTTCGGCAGTTCGAAATCGTGGCGGCTCGACAGCGGCAACGCA 5820  
Db 5761 GAGTCCGCCACAAAGACCTTCGGCAGTTCGAAATCGTGGCGGCTCGACAGCGGCAACGCA 5820  
QY 5821 ACGGCTCTCTGACAGCCCTCCGACGACGCGGATCCGATCTCAGCGAGGCTGAGTCTGATC 5880  
Db 5821 ACGGCTCTCTGACAGCCCTCCGACGACGCGGATCCGATCTCAGCGAGGCTGAGTCTGATC 5880  
QY 5881 TCCTTCCATGCCCCCTTGGGGGAGCGCGGGGATCCGATCTCAGCGAGGCTGTTGG 5940  
Db 5881 TCCTTCCATGCCCCCTTGGGGGAGCGCGGGGATCCGATCTCAGCGAGGCTGTTGG 5940  
QY 5941 TCTACCGTAGCGAGGAGGCTAGTGGAGAGTGTCTGTCTGTCTGATGTCTTACATGG 6000  
Db 5941 TCTACCGTAGCGAGGAGGCTAGTGGAGAGTGTCTGTCTGTCTGATGTCTTACATGG 6000  
QY 6001 ACAGCGCCCTGATCACGCCATGCTGCGGAGGAAACCAAGCTGCCCATCAATGCACTG 6060  
Db 6001 ACAGCGCCCTGATCACGCCATGCTGCGGAGGAAACCAAGCTGCCCATCAATGCACTG 6060  
QY 6061 AGCAACTCTTTGCTCCGTCACCAACCTTGGTCTATGCTACAAACATCTCGCAGCGCAAGC 6120

6061	Db	AGCAACTCTTTGTGCTCCGTCAACCAACTTGTGCTATGCTACAAATCTCGACGGCAAGC	6120
6121	Qy	CTCGGGCAGAGAAGGTCACCTTTGACAGACTCGAGTCTCTGACGACCACTACCCGGGAC	6180
6121	Db	CTCGGGCAGAGAAGGTCACCTTTGACAGACTCGAGTCTCTGACGACCACTACCCGGGAC	6180
6181	Qy	GTGCTCAAGGAGATGAAGGGGAAGGCGTCACACAGTTAAAGGCTAAATCTTATCCGTGGAG	6240
6181	Db	GTGCTCAAGGAGATGAAGGGGAAGGCGTCACACAGTTAAAGGCTAAATCTTATCCGTGGAG	6240
6241	Qy	GAAGCCTGTAAGCTGAGCCGCCACATTCGGCCAGATCTAAATTTGGCTATGGGGCAAG	6300
6241	Db	GAAGCCTGTAAGCTGAGCCGCCACATTCGGCCAGATCTAAATTTGGCTATGGGGCAAG	6300
6301	Qy	GACGTCGGAACTTATCCAGCAAGCGCGTTAAACACATCCGCTCCGTGTGGGAAGGACTTG	6360
6301	Db	GACGTCGGAACTTATCCAGCAAGCGCGTTAAACACATCCGCTCCGTGTGGGAAGGACTTG	6360
6361	Qy	CTGGAAGACACTGAGACACCAATTTGACACCACTCATATGGCAAAAATAGAGTTTTCTGC	6420
6361	Db	CTGGAAGACACTGAGACACCAATTTGACACCACTCATATGGCAAAAATAGAGTTTTCTGC	6420
6421	Qy	GTCCAAACGAGAAAGGGGGCGCAAGCCAGTCGCTTATCGTATTTCCAGATTTTGGGG	6480
6421	Db	GTCCAAACGAGAAAGGGGGCGCAAGCCAGTCGCTTATCGTATTTCCAGATTTTGGGG	6480
6481	Qy	GTTTCGTGTGTCGAGAAATAGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCCGTG	6540
6481	Db	GTTTCGTGTGTCGAGAAATAGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCCGTG	6540
6541	Qy	ATGGGCTCTTCATACGGATTCCAAATCTCTCTGGACAGCGGGTCGAGTTCTCTGGTGAAT	6600
6541	Db	ATGGGCTCTTCATACGGATTCCAAATCTCTCTGGACAGCGGGTCGAGTTCTCTGGTGAAT	6600
6601	Qy	GCCTGGAAGCGAAGAAATGCGCTTATGGGCTTCGCATATGACACCCGCTGTTTGACTCA	6660
6601	Db	GCCTGGAAGCGAAGAAATGCGCTTATGGGCTTCGCATATGACACCCGCTGTTTGACTCA	6660
6661	Qy	ACGCTCACTGAGATGATCANTCCGTGTTGAGAGTCAATCTACCAATGCTGTGACTGGCC	6720
6661	Db	ACGCTCACTGAGATGATCANTCCGTGTTGAGAGTCAATCTACCAATGCTGTGACTGGCC	6720
6721	Qy	CCCGAAGCCAGACAGGCCATTAAGTTCGCTCACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780
6721	Db	CCCGAAGCCAGACAGGCCATTAAGTTCGCTCACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780
6781	Qy	ACTAATTTCTAAGGGCGAACTGCGGCTATGCGCGGTCCGCGGAGCGGTGTACTGACG	6840
6781	Db	ACTAATTTCTAAGGGCGAACTGCGGCTATGCGCGGTCCGCGGAGCGGTGTACTGACG	6840
6841	Qy	ACCAAGCTCGGTAATACCTTCACATGTTACTTTGAAGGCCGCTCGCGGCTCTCGAGCTGG	6900
6841	Db	ACCAAGCTCGGTAATACCTTCACATGTTACTTTGAAGGCCGCTCGCGGCTCTCGAGCTGG	6900
6901	Qy	AAGCTCCAGGACTCGACGATCTCGTATGCGGAGACGACCTTGTGCTTATCTGTGAAAGC	6960
6901	Db	AAGCTCCAGGACTCGACGATCTCGTATGCGGAGACGACCTTGTGCTTATCTGTGAAAGC	6960
6961	Qy	GCGGGGACCCAAAGAGGACGAGCGAGGCTACGGGCTTCAACAGAAATGAGTTGATAATCATGC	7020
6961	Db	GCGGGGACCCAAAGAGGACGAGCGAGGCTACGGGCTTCAACAGAAATGAGTTGATAATCATGC	7020
7021	Qy	TCTGCCCCCTCTGGGACCCGCCAACCAAGATACGACTTGGAGTTGATAATCATATGC	7080
7021	Db	TCTGCCCCCTCTGGGACCCGCCAACCAAGATACGACTTGGAGTTGATAATCATATGC	7080
7081	Qy	TCCTCCAATGTCTAGTCGGCGACGATGCACTCTGGCAAAAGGGTGTACTATCTCACCCGT	7140
7081	Db	TCCTCCAATGTCTAGTCGGCGACGATGCACTCTGGCAAAAGGGTGTACTATCTCACCCGT	7140
7141	Qy	GACCCCAACACCCCTCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT	7200

Db	7141	GACCCACACACCCCTTGGCGGGCTGCTGGGAGACAGTAGACACACTCCAGTCAAT	7200
Qy	7201	TCCTGGCTAGGCAACATCATATGATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
Db	7201	TCCTGGCTAGGCAACATCATATGATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
Qy	7261	ACTCATTTCTTCTCCATCTCTTAGCTCAGGAACAACCTTGAAAAGCCCTAGATTGTCTAG	7320
Db	7261	ACTCATTTCTTCTCCATCTCTTAGCTCAGGAACAACCTTGAAAAGCCCTAGATTGTCTAG	7320
Qy	7321	ATCTACGGGGCTGTACTCCATTTAGCCACTTTGACCTACTCTCAGATCATTTCAACGACTC	7380
Db	7321	ATCTACGGGGCTGTACTCCATTTAGCCACTTTGACCTACTCTCAGATCATTTCAACGACTC	7380
Qy	7381	CATGGCCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT	7440
Db	7381	CATGGCCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT	7440
Qy	7441	TCATGCTCAGGAAACTTTGGGGTACCGCCCTTGGGAGTCTGGAGACATCGGGCCAGAAGT	7500
Db	7441	TCATGCTCAGGAAACTTTGGGGTACCGCCCTTGGGAGTCTGGAGACATCGGGCCAGAAGT	7500
Qy	7501	GTCGCGCTAGGCTACTGTGCCAGGGGGAGGGCTGCCACTTCTGCGCAGTACTCTCTTC	7560
Db	7501	GTCGCGCTAGGCTACTGTGCCAGGGGGAGGGCTGCCACTTCTGCGCAGTACTCTCTTC	7560
Qy	7561	AACTGGGCAGTAGAGCAACAAGCTCAAACTCACTCCAATCCCGGCTGGTCCACGTTGGAT	7620
Db	7561	AACTGGGCAGTAGAGCAACAAGCTCAAACTCACTCCAATCCCGGCTGGTCCACGTTGGAT	7620
Qy	7621	TTATCCAGCTGGTTGCTTGTGTTACAGCGGGGAGACATATATCACAGCTGTCTCGT	7680
Db	7621	TTATCCAGCTGGTTGCTTGTGTTACAGCGGGGAGACATATATCACAGCTGTCTCGT	7680
Qy	7681	GCCGACCCCGCTGGTTCATGTGTCGCTACTCCTACTCTTCTGTAGGGGTAGGCATCTAT	7740
Db	7681	GCCGACCCCGCTGGTTCATGTGTCGCTACTCCTACTCTTCTGTAGGGGTAGGCATCTAT	7740
Qy	7741	CTACTCCCCAAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT	7800
Db	7741	CTACTCCCCAAACCGATGAACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT	7800
Qy	7801	TTTCCCTTTTTTTTTTCTT	7860
Db	7801	TTTCCCTTTTTTTTTTCTT	7860
Qy	7861	TTTTTTCCTCTTTTTTTCCTTTTCTTCTTGTGGTGGCTCCATCTTAGCCCTAGTCACGGC	7920
Db	7861	TTTTTTCCTCTTTTTTTCCTTTTCTTCTTGTGGTGGCTCCATCTTAGCCCTAGTCACGGC	7920
Qy	7921	TAGCTGTGAAGGTCCTGTAGCCCTTTGACTGACAGAGAGTGTCTGATACTTGGCCTCTCTGC	7980
Db	7921	TAGCTGTGAAGGTCCTGTAGCCCTTTGACTGACAGAGAGTGTCTGATACTTGGCCTCTCTGC	7980
Qy	7981	AGATCAAGT 7989	
Db	7981	AGATCAAGT 7989	
RESULT 5			
AR406042			
LOCUS	AR406042	8001 bp	linear
DEFINITION	Sequence 7 from patent US 6630343.		
ACCESSION	AR406042		
VERSION	AR406042.1	GI:40155169	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 8001)		
AUTHORS	Bartenschlager, R.		
TITLE	Hepatitis C virus culture system		
JOURNAL	Patent: US 6630343-A 7 07-OCT-2003;		



2029 QY ACCAATGTGGACACAGAACTCTGTGGCTGGGAAAGCGCCCCCGGGGCGGTTCTCTTGACA 2088  
2041 Db ACCAATGTGGACACAGAACTCTGTGGCTGGGAAAGCGCCCCCGGGGCGGTTCTCTTGACA 2100  
2089 QY CCATGCACCTCGGGAGCTCGGACCTTTACTTTGGTCACGAGCATGSCCGATGTCATTCCG 2148  
2101 Db CCATGCACCTCGGGAGCTCGGACCTTTACTTTAGTTCAGAGGATGCCGATGTCAITTCG 2160  
2149 QY GTGCGCGCGGGGGGACACAGCAGGGGAGCCTACTCTCCCCCAGCGCCGCTCTCTACTTG 2208  
2161 Db GTGCGCGCGGGGACACAGCAGGGGAGCCTACTCTCCCCCAGCGCCGCTCTCTACTTG 2220  
2209 QY AAGGGCTCTTTCGGGCGGTCCACTGCTCTGCCCTCGGGGCACTGTGTGGGCATCTTTCG 2268  
2221 Db AAGGGCTCTTTCGGGCGGTCCACTGCTCTGCCCTCGGGGCACTGTGTGGGCATCTTTCG 2280  
2269 QY GCTGCCGTGTGACCCGAGGGGTTCGAAAGCGGTGGACTTTGTACCCGTCGAGTCTATG 2328  
2281 Db GCTGCCGTGTGACCCGAGGGGTTCGAAAGCGGTGGACTTTGTACCCGTCGAGTCTATG 2340  
2329 QY GAAACCACTATGCGGTCCCGGTCTTTCAGGCAACTCGTCCCTCCGSCCGTACCGCAG 2388  
2341 Db GAAACCACTATGCGGTCCCGGTCTTTCAGGCAACTCGTCCCTCCGSCCGTACCGCAG 2400  
2389 QY ACATTCAGGTGCCCATCTACAGCCCTCTACTGTGTAGCGGCAAGACATAAGTGGCG 2448  
2401 Db ACATTCAGGTGCCCATCTACAGCCCTCTACTGTGTAGCGGCAAGACATAAGTGGCG 2460  
2449 QY GCTGCCGTGTGACCCGAGGGGTTCGAAAGCGGTGGACTTTGTACCCGTCGAGTCTATG 2508  
2461 Db GCTGCCGTGTGACCCGAGGGGTTCGAAAGCGGTGGACTTTGTACCCGTCGAGTCTATG 2520  
2509 QY CTAGTTCCTGGGCGGTATGCTTAAGGCACATGCTATCGACCCCTAACATCAGAACCGGG 2568  
2521 Db CTAGTTCCTGGGCGGTATGCTTAAGGCACATGCTATCGACCCCTAACATCAGAACCGGG 2580  
2569 QY GTAAGACCATACACAGGGTGGCCCCCATCAGTACTCCACTATGGCAAGTTCCTG 2628  
2581 Db GTAAGACCATACACAGGGTGGCCCCCATCAGTACTCCACTATGGCAAGTTCCTG 2640  
2629 QY GACGGTGTGCTCTGGGCGGCTATGACATCATATATGTATGATGCTCCACTCACT 2688  
2641 Db GACGGTGTGCTCTGGGCGGCTATGACATCATATATGTATGATGATGCTCCACTCACT 2700  
2689 QY GACTCGACCACTATCTCTGGGCTCGGCACAGTCTCGAACAGCGGAGCGGTGGAGCG 2748  
2701 Db GACTCGACCACTATCTCTGGGCTCGGCACAGTCTCGAACAGCGGAGCGGTGGAGCG 2760  
2749 QY CGACTGCTGTGCTCGGCACCGCTACGCCCTCGGGATCGGTCAACGTCACATCCAAAC 2808  
2761 Db CGACTGCTGTGCTCGGCACCGCTACGCCCTCGGGATCGGTCAACGTCACATCCAAAC 2820  
2809 QY ATCGAGAGGTGCTCTGTCCAGCAGTGGAGAAATCCCTTTTATGCCAAAGCCATCC 2868  
2821 Db ATCGAGAGGTGCTCTGTCCAGCAGTGGAGAAATCCCTTTTATGCCAAAGCCATCC 2880  
2869 QY ATCGAGACCATCAAGGGGGGAGGCACCTCATTTTCTGCCATTCCAAAGAAATGTGAT 2928  
2881 Db ATCGAGACCATCAAGGGGGGAGGCACCTCATTTTCTGCCATTCCAAAGAAATGTGAT 2940  
2929 QY GAGTCCCGGGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTT 2988  
2941 Db GAGTCCCGGGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTT 3000  
2989 QY GATGTATTCGTCATACCACATAGCGGAGAGCTCATTTGTGTAGCAACCGAGCTCTAATG 3048  
3001 Db GATGTATTCGTCATACCACATAGCGGAGAGCTCATTTGTGTAGCAACCGAGCTCTAATG 3060  
3049 QY ACGGGCTTTACCGGCGATTTTCGACTCAGTGTATCGACTGCAATACATGTGACCCAGACA 3108  
3061 Db ACGGGCTTTACCGGCGATTTTCGACTCAGTGTATCGACTGCAATACATGTGACCCAGACA 3120  
3109 QY GTCGACTTCAGCTCGGACCCGACCTTCACCATTTGAGACGAGCCGTGCGCAAGAGCGG 3168

3121 Db GTGCGACTTCAGCTCGAACCCGACCTTCACATTTGAGACGACGACCTGCGCAACAAGCGG 3180  
3169 QY GTGTCACTGCTCGACGCGGAGCAGGACTGTGTAGGGGAGATGGGCATTTTACAGTTT 3228  
3181 Db GTGTCACTGCTCGACGCGGAGCAGGACTGTGTAGGGGAGATGGGCATTTTACAGTTT 3240  
3229 QY GTGACTCCAGGAGAACGGCCCTCGGGCATGTTTCGATTCCTCGGTTCTGTGCGAGTCTAT 3288  
3241 Db GTGACTCCAGGAGAACGGCCCTCGGGCATGTTTCGATTCCTCGGTTCTGTGCGAGTCTAT 3300  
3289 QY GACGGGGCTGTGCTTGTGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTTGGCGGT 3348  
3301 Db GACGGGGCTGTGCTTGTGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTTGGCGGT 3360  
3349 QY TACTTAAACACACCAAGGTTGCCGCTGCCAGGACCATCTGGAGTCTGGGAGAGCGTC 3408  
3361 Db TACTTAAACACACCAAGGTTGCCGCTGCCAGGACCATCTGGAGTCTGGGAGAGCGTC 3420  
3409 QY TTTTACAGGCTCACCCACATAGACGCCCATTTTCTTGTCCAGACTAAGCAGGCGAGAGAC 3468  
3421 Db TTTTACAGGCTCACCCACATAGACGCCCATTTTCTTGTCCAGACTAAGCAGGCGAGAGAC 3480  
3469 QY AACTTCCCCTACTCTGTAGCATACAGGCTACGGTGTGCCAGGCTCAGGCTCCACCT 3528  
3481 Db AACTTCCCCTACTCTGTAGCATACAGGCTACGGTGTGCCAGGCTCAGGCTCCACCT 3540  
3529 QY CCATCTGGGACCAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACGCGGCCA 3588  
3541 Db CCATCTGGGACCAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACGCGGCCA 3600  
3589 QY AGCCCCCTCTGTATAGGCTGGGAGCCGTTTCAAAAGAGGTTACTACACACACCCCAT 3648  
3601 Db AGCCCCCTCTGTATAGGCTGGGAGCCGTTTCAAAAGAGGTTACTACACACACCCCAT 3660  
3649 QY ACCAAATACATCATGTCATGTCGCTGACCTCGAGGTCGTCACGAGCACTGGGTG 3708  
3661 Db ACCAAATACATCATGTCATGTCGCTGACCTCGAGGTCGTCACGAGCACTGGGTG 3720  
3709 QY CTGTTAGGCGGAGTCTAGCAGCTGTGGCGGATGTCCTGTGACAAAGCAGCGTGGTC 3768  
3721 Db CTGTTAGGCGGAGTCTAGCAGCTGTGGCGGATGTCCTGTGACAAAGCAGCGTGGTC 3780  
3769 QY ATTGTGGGAGGATCATCTGTCCGAAAGCCGGCCATCATTTCCGACAGGGAAGTCTT 3828  
3781 Db ATTGTGGGAGGATCATCTGTCCGAAAGCCGGCCATCATTTCCGACAGGGAAGTCTT 3840  
3829 QY TACCGGAGTTCGATGAGATGGAAGAGTGGCTCACACCTCCCTTACATCGAACAGGGA 3888  
3841 Db TACCGGAGTTCGATGAGATGGAAGAGTGGCTCACACCTCCCTTACATCGAACAGGGA 3900  
3889 QY ATGAGCTCGCCGAAACAAATTCAAACAGAGGCAATCGGTTGTGCAAAAGCAGCCCAAG 3948  
3901 Db ATGAGCTCGCCGAAACAAATTCAAACAGAGGCAATCGGTTGTGCAAAAGCAGCCCAAG 3960  
3949 QY CAAGCGGAGCTGCTGCTCCGTTGGTGAATCCAGTGGGAGCCCTCGAGGCTTCTG 4008  
3961 Db CAAGCGGAGCTGCTGCTCCGTTGGTGAATCCAGTGGGAGCCCTCGAGGCTTCTG 4020  
4009 QY GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTG 4068  
4021 Db GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTG 4080  
4069 QY CCTGGCAACCCCGCATAGCATCACTGATGGCATTCACGCTCTATFACAGCCGCTC 4128  
4081 Db CCTGGCAACCCCGCATAGCATCACTGATGGCATTCACGCTCTATFACAGCCGCTC 4140  
4129 QY ACCACCCAAACATACCTCTCTTTTAACTCTCGGGGATGGGTGGCCGCCCACTTGT 4188  
4141 Db ACCACCCAAACATACCTCTCTCTTTTAACTCTCGGGGATGGGTGGCCGCCCACTTGT 4200  
4189 QY CCTCCAGCGCTCTCTCTCTCTTTCTGTTAGCGCGCGGATCGCTGAGCGGCTGTGGCAGC 4248

Db	4201	CCTCCGACGCTGCTTCTGCTTTTGGTAGCGCCGGCAGTCGCTGGAGCGCGCTGTTGGCAGC	4261
QY	4249	ATAGSCCTTGGGAAGGTGCTTTGTGGATATTTTGGCAGGTTATGAGCAGAGGGTGGCAGGC	4308
Db	4261	ATAGSCCTTGGGAAGGTGCTTTGTGGATATTTTGGCAGGTTATGAGCAGAGGGTGGCAGGC	4320
QY	4309	GCGCTGTCGGCTTTAAGTTCATGAGCGGCGAGATGCCTCCACCGAGGACCTGGTTAAC	4368
Db	4321	GCGCTGTCGGCTTTAAGTTCATGAGCGGCGAGATGCCTCCACCGAGGACCTGGTTAAC	4380
QY	4369	CTACTCCCTGCTATCCTCTCCCTTGGCGCCCTAGTCTGTCGGGTGCTGTGGCAGCGATA	4428
Db	4381	CTACTCCCTGCTATCCTCTCCCTTGGCGCCCTAGTCTGTCGGGTGCTGTGGCAGCGATA	4440
QY	4429	CTGGCTCGGCACGTGGGCCACAGGGAGGGGGCTGTGCAGTGGATGAACCCGGCTGATAGCG	4488
Db	4441	CTGGCTCGGCACGTGGGCCACAGGGAGGGGGCTGTGCAGTGGATGAACCCGGCTGATAGCG	4500
QY	4489	TTCCGTTTCGGGGGTAAACACGTCTCCCCACGCACCTATGTGCTTGAGACGACGCTGCAC	4548
Db	4501	TTCCGTTTCGGGGGTAAACACGTCTCCCCACGCACCTATGTGCTTGAGACGACGCTGCAC	4560
QY	4549	GCAGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGCCTTCAC	4608
Db	4561	GCAGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGCCTTCAC	4620
QY	4609	CAGTGGATCAACGAGGACTGCTCCACGCCATGCTCCGGCTCTGGGCTTAAGAGATGTTGG	4668
Db	4621	CAGTGGATCAACGAGGACTGCTCCACGCCATGCTCCGGCTCTGGGCTTAAGAGATGTTGG	4680
QY	4669	GATTGGATATGCACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCCAAGCTCTGCCG	4728
Db	4681	GATTGGATATGCACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCCAAGCTCTGCCG	4740
QY	4729	CGATTTGCGGGAGTCCCTTCTCTCATGTACAGTGGGTAAAGGGAGTCTGGCGGGC	4788
Db	4741	CGATTTGCGGGAGTCCCTTCTCTCATGTACAGTGGGTAAAGGGAGTCTGGCGGGC	4800
QY	4789	GACGGCATCATGCAAAACACCTGCCATGTGGAGCACAGATCACCGGACATGTGMAAAC	4848
Db	4801	GACGGCATCATGCAAAACACCTGCCATGTGGAGCACAGATCACCGGACATGTGMAAAC	4860
QY	4849	GTTTCCATAGGATCTGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGAACTTCCC	4908
Db	4861	GTTTCCATAGGATCTGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGAACTTCCC	4920
QY	4909	ATTAAACGCTACACCGGGCCCTTCAGCCCTCCCGGGCGCAAAATTATTCAGGGCG	4968
Db	4921	ATTAAACGCTACACCGGGCCCTTCAGCCCTCCCGGGCGCAAAATTATTCAGGGCG	4980
QY	4969	CTGTGGCGGTGCTGTGAGGAGTACGTGCAGGTTACGCGGTGGGGATTTCACATAC	5028
Db	4981	CTGTGGCGGTGCTGTGAGGAGTACGTGCAGGTTACGCGGTGGGGATTTCACATAC	5040
QY	5029	GTACGGGCATGACCACTGACAAAGTGGCCGTGTCAAGTTCGGGCCCCCGAATTC	5088
Db	5041	GTACGGGCATGACCACTGACAAAGTGGCCGTGTCAAGTTCGGGCCCCCGAATTC	5100
QY	5089	TTCAACAAGTGGATGGGTGCGGTGACAGGTACGCTCCAGCGTGCAAAACCCCTCTTA	5148
Db	5101	TTCAACAAGTGGATGGGTGCGGTGACAGGTACGCTCCAGCGTGCAAAACCCCTCTTA	5160
QY	5149	CGGGAGAGGTACATTCCTTGGTCGGGCTCAATCAATACCTGGTTGGGTACAGCTCCCA	5208
Db	5161	CGGGAGAGGTACATTCCTTGGTCGGGCTCAATCAATACCTGGTTGGGTACAGCTCCCA	5220
QY	5209	TGGAGCGCGAACCGGACGTAGCAGTGTCACTTCCATGCTCACCGACCCCTCCACATTC	5268
Db	5221	TGGAGCGCGAACCGGACGTAGCAGTGTCACTTCCATGCTCACCGACCCCTCCACATTC	5280
QY	5269	ACGGCGGACAGCCTTAAGCTAGCTGGCCAGGGGATCTCCCTCTCTTTGGCCAGCTCA	5328
Db	5281	ACGGCGGACAGCCTTAAGCTAGCTGGCCAGGGGATCTCCCTCTCTTTGGCCAGCTCA	5340



6409 QY GAGTTTCTGCGTCCAAACGAGAGAGGGGGCCGCAAGCCAGCTCGCCTTATCGTATTC 6468  
6421 Db GAGGTTTCTGCGTCCAAACGAGAGAGGGGGCCGCAAGCCAGCTCGCCTTATCGTATTC 6480  
6469 QY CCAGATTTGGGGGTTGCTGTGCGCAGAAAATGGCCCTTACGATGTGCTCTCCACCCCTC 6528  
6481 Db CCAGATTTGGGGGTTGCTGTGCGCAGAAAATGGCCCTTACGATGTGCTCTCCACCCCTC 6540  
6529 QY CCTCAGGCGCTGATGGGCTCTTATACGGAATCCAAATCTCTCTGGACAGCGGTTCGAG 6588  
6541 Db CCTCAGGCGCTGATGGGCTCTTATACGGAATCCAAATCTCTCTGGACAGCGGTTCGAG 6600  
6589 QY TTCCTGGTGAATGCTGTGAAACGAGAAAATGCCCTATGCGCTTCGCATATGACACCCGC 6648  
6601 Db TTCCTGGTGAATGCTGTGAAACGAGAAAATGCCCTATGCGCTTCGCATATGACACCCGC 6660  
6649 QY TGTTTTGACTCAACGGTCACTCAGAAATGATCCGCTGTGAGGAGTCAATCTACCAATGT 6708  
6661 Db TGTTTTGACTCAACGGTCACTCAGAAATGATCCGCTGTGAGGAGTCAATCTACCAATGT 6720  
6709 QY TGTGACTTGGCCCCGGAAGCCAGACAGGCCATAAGTCCGCTCAAGAGCGGCTTTACATC 6768  
6721 Db TGTGACTTGGCCCCGGAAGCCAGACAGGCCATAAGTCCGCTCAAGAGCGGCTTTACATC 6780  
6769 QY GGGGGCCCCCTGACTAAATTTCTAAAGGGCAGAACTGCGGCTATGCGCGGTGCGCGCGAGC 6828  
6781 Db GGGGGCCCCCTGACTAAATTTCTAAAGGGCAGAACTGCGGCTATGCGCGGTGCGCGCGAGC 6840  
6829 QY GGTGTACTCAGCAGCAGCTGCGGTATACCCCTACATGTTACTTTGAAGCCGCTCGCGGC 6888  
6841 Db GGTGTACTCAGCAGCAGCTGCGGTATACCCCTACATGTTACTTTGAAGCCGCTCGCGGC 6900  
6889 QY TGTGAGCTCGGAAGCTCCAGAGCTCAGCATGCTCGTATGCGGAGACGACCTTGTGTT 6948  
6901 Db TGTGAGCTCGGAAGCTCCAGAGCTCAGCATGCTCGTATGCGGAGACGACCTTGTGTT 6960  
6949 QY ATCTGTGAAGCGGGGACCCNAGAGGAGGAGCGAGCTACGGCCTTCAGGAGGCT 7008  
6961 Db ATCTGTGAAGCGGGGACCCNAGAGGAGGAGCGAGCTACGGCCTTCAGGAGGCT 7020  
7009 QY ATGACTAGATATCTGCCCCCTCGGGGACCCGCAACCAAGATACGACTTGGAGTTG 7068  
7021 Db ATGACTAGATATCTGCCCCCTCGGGGACCCGCAACCAAGATACGACTTGGAGTTG 7080  
7069 QY ATAAACATCATGCTCTCCAAATGTGTGTCAGTCGCGCAGCATCTCGGCAAAAGGGTGTAC 7128  
7081 Db ATAAACATCATGCTCTCCAAATGTGTGTCAGTCGCGCAGCATCTCGGCAAAAGGGTGTAC 7140  
7129 QY TATCTCAGCGTGACCCGACACCCCTTGGCGGGCTGCGTGGGAGACAGTACAC 7188  
7141 Db TATCTCAGCGTGACCCGACACCCCTTGGCGGGCTGCGTGGGAGACAGTACAC 7200  
7189 QY ACTCCAGTCAATTCCTGGCTAGGCAACATCATGATGTCGCCCACTTGTGGGCAAGG 7248  
7201 Db ACTCCAGTCAATTCCTGGCTAGGCAACATCATGATGTCGCCCACTTGTGGGCAAGG 7260  
7249 QY ATGATCCTGATGACTCATTTCTTCTCCATCCTTCTAGCTCAGGAACAACCTTGAAAAGCC 7308  
7261 Db ATGATCCTGATGACTCATTTCTTCTCCATCCTTCTAGCTCAGGAACAACCTTGAAAAGCC 7320  
7309 QY CTAGATGTCAGATCTACGGGGCTGTACTCATTGAGCCACTTGACCTACCTCAGATC 7368  
7321 Db CTAGATGTCAGATCTACGGGGCTGTACTCATTGAGCCACTTGACCTACCTCAGATC 7380  
7369 QY ATTCAAGCACTCCATGSCCTTAGCGCAATTTCACTCCATAGTTACTCTCAGGTGAGATC 7428  
7381 Db ATTCAAGCACTCCATGSCCTTAGCGCAATTTCACTCCATAGTTACTCTCAGGTGAGATC 7440  
7429 QY AATAGGTTGGCTTCATGCCCTCAGGAACCTTGGGTACCGCCCTTGGAGTCTGGAGACAT 7488  
7441 Db AATAGGTTGGCTTCATGCCCTCAGGAACCTTGGGTACCGCCCTTGGAGTCTGGAGACAT 7500  
7489 QY CGGCGCAGAAAGTGTCCCGCTAGGCTACTGTGCCAGGGGGGGGCGTCCACTTGTGGC 7548

7501 Db CGGCGCAGAAAGTGTCCGGCTAGGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGGC 7560  
7549 QY AAGTACCTCTTTCAACTGGCGAGTAAGGACCAAGCTCAAACTCAGTCCAAATCCCGGCTGG 7608  
7561 Db AAGTACCTCTTTCAACTGGCGAGTAAGGACCAAGCTCAAACTCAGTCCAAATCCCGGCTGG 7620  
7609 QY TCCAGTTGGATTTATCCAGCTGGTTCGTTGCTGGTTACAGGGGGGAGACATATAC 7668  
7621 Db TCCAGTTGGATTTATCCAGCTGGTTCGTTGCTGGTTACAGGGGGGAGACATATAC 7680  
7669 QY AGCTGTCTCGTCCCGGACCCCGCTGGTTCAATGTGGTCTACTCTACTTCTGTAGGG 7728  
7681 Db AGCTGTCTCGTCCCGGACCCCGCTGGTTCAATGTGGTCTACTCTACTTCTGTAGGG 7740  
7729 QY GFAGCATCTATCTACTCTCCCAACCGATGAACGGGACCTAAACACTCCAGCCCAATAG 7788  
7741 Db GFAGCATCTATCTACTCTCCCAACCGATGAACGGGAGCTAAACACTCCAGCCCAATAG 7800  
7789 QY CCATCCTGTTTTTTTCCCTTT 7848  
7801 Db CCATCCTGTTTTTTTCCCTTT 7860  
7849 QY TTTTCTCCTTTTTTTTTTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7908  
7861 Db TTTTCTCCTTTTTTTTTTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7920  
7909 QY CCTAGTCACGGTAGCTGTGAAGGTCGTGAGCGCTGTGACTGCAGAGAGTGTGATAC 7968  
7921 Db CCTAGTCACGGTAGCTGTGAAGGTCGTGAGCGCTGTGACTGCAGAGAGTGTGATAC 7980  
7969 QY TGGCCTCTCTGCAGATCAAGT 7989  
7981 Db TGGCCTCTCTGCAGATCAAGT 8001

RESULT 6  
AX036254  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
ORIGIN

AX036254  
Sequence 3 from Patent EP1043399.  
AX036254  
AX036254.1 GI:11225870  
Hepatitis C virus  
Hepatitis C virus  
viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
Hepacivirus.  
1  
Bartenschlager, R.D.  
Hepatitis c virus cell culture system  
Patent: EP 1043399-A 3 11-OCT-2000;  
BARTENSCHLAGER RALF DR (DE)  
Location/Qualifiers  
1. .8001  
/organism="Hepatitis C virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11103"

Query Match 99.7%; Score 7965.4; DB 6; Length 8001;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 7989; Conservative 0; Mismatches 1; Indels 12; Gaps 1;  
QY 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAAGTACTG 60  
Db 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAAGTACTG 60  
QY 61 TCCTCAGCAGAAAGCGTCTAGCCATGGCTTAGTATGATGTCGTCGAGCTCCAGGAC 120  
Db 61 TCCTCAGCAGAAAGCGTCTAGCCATGGCTTAGTATGATGTCGTCGAGCTCCAGGAC 120  
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGGAAACGGGTGAGTACACCGGAATTCGAC 180



QY 2329 GAAACCACTATGCGGTCCCGGTCTTACGGACAACCTGGTCCCTCCGCGGTACCGGAC 2388  
DB 2341 GAAACCACTATGCGGTCCCGGTCTTACGGACAACCTGGTCCCTCCGCGGTACCGGAC 2400  
QY 2389 ACATTTCCAGTGGCCCATCTACAGCCCTCTCTGTTAGCGGCAAGAGCACTAAGTGGCG 2448  
DB 2401 ACATTTCCAGTGGCCCATCTACAGCCCTCTCTGTTAGCGGCAAGAGCACTAAGTGGCG 2460  
QY 2449 GCTGGTATGACGCCCAAGGGTATAAGGTGCTTGTCTGAAACCCGTCCTGCGCCGACCC 2508  
DB 2461 GCTGGTATGACGCCCAAGGGTATAAGGTGCTTGTCTGAAACCCGTCCTGCGCCGACCC 2520  
QY 2509 CTAGGTTTCGGGGCGTATATGTTCTAAGGCACATGTTATCGACCTTAAACATGACAGACCGG 2568  
DB 2521 CTAGGTTTCGGGGCGTATATGTTCTAAGGCACATGTTATCGACCTTAAACATGACAGACCGG 2580  
QY 2569 GTAAGGACCATCACAGCGGTGCCCCCATCAGTACTCTCCACTATGCGGAAGTTTCTTGCC 2628  
DB 2581 GTAAGGACCATCACAGCGGTGCCCCCATCAGTACTCTCCACTATGCGGAAGTTTCTTGCC 2640  
QY 2629 GACGGTGTGTTGCTCTGGGGCGCTATGACATCATATATGATGATGAGTGCCACTCACT 2688  
DB 2641 GACGGTGTGTTGCTCTGGGGCGCTATGACATCATATATGATGATGAGTGCCACTCACT 2700  
QY 2689 GACTCGACCACTATCTCTGGGCAATCGGCACAGTCTCTGGACAAGCGGAGCGTGGAGCG 2748  
DB 2701 GACTCGACCACTATCTCTGGGCAATCGGCACAGTCTCTGGACAAGCGGAGCGTGGAGCG 2760  
QY 2749 CCACTCTGTCGTCTGCCACCGCTAGCCTCCGGGATCGGTCACTGTCGACATCCAAAC 2808  
DB 2761 CCACTCTGTCGTCTGCCACCGCTAGCCTCCGGGATCGGTCACTGTCGACATCCAAAC 2820  
QY 2809 ATCGAGAGGTGGTCTGTCTGACACTGAGAAATCCCTTTATGGCAAGCAATCCCT 2868  
DB 2821 ATCGAGAGGTGGTCTGTCTGACACTGAGAAATCCCTTTATGGCAAGCAATCCCT 2880  
QY 2869 ATCGAGACCATCAAGGGGGAGGACCTCACTTTCTGCGATTCGAAGAAGTGTGAT 2928  
DB 2881 ATCGAGACCATCAAGGGGGAGGACCTCACTTTCTGCGATTCGAAGAAGTGTGAT 2940  
QY 2929 GAGTTCGCGGGAAGCTGTCGGGCTCGGACTCAATGCTGTAGCATATTACGGGGCTTT 2988  
DB 2941 GAGTTCGCGGGAAGCTGTCGGGCTCGGACTCAATGCTGTAGCATATTACGGGGCTTT 3000  
QY 2989 GATGTATCGGTATACCAACTAGCGGAGAGCTGATTTGCTAGCAACGAGCGCTCTAATG 3048  
DB 3001 GATGTATCGGTATACCAACTAGCGGAGAGCTGATTTGCTAGCAACGAGCGCTCTAATG 3060  
QY 3049 ACGGGCTTTACGGCGAATTCGACTCAGTGATCGACTGCAATACATGTGTCAACCGAGACA 3108  
DB 3061 ACGGGCTTTACGGCGAATTCGACTCAGTGATCGACTGCAATACATGTGTGTCAACCGAGACA 3120  
QY 3109 GTGACTTCAGCTGAGCCGACCTTACCATTTGAGACGACGACCGTGCACAAAGACGG 3168  
DB 3121 GTGACTTCAGCTGAGCCGACCTTACCATTTGAGACGACGACCGTGCACAAAGACGG 3180  
QY 3169 GTGTACGCTTCGAGCGGAGGAGGAGCTGTTAGGGGCAAGATGGGCATTTACAGTTT 3228  
DB 3181 GTGTACGCTTCGAGCGGAGGAGGAGCTGTTAGGGGCAAGATGGGCATTTACAGTTT 3240  
QY 3229 GTGACTTCAGGAGAAAGGGCCCTCGGGATGTTGATTCCTGCTGCGAGTGTAT 3288  
DB 3241 GTGACTTCAGGAGAAAGGGCCCTCGGGATGTTGATTCCTGCTGCGAGTGTAT 3300  
QY 3289 GACGGGGCTGCTGTTGATACGAGCTCACCGCGCGAGACCTCAGTTAGTTGTTGGGGCT 3348  
DB 3301 GACGGGGCTGCTGTTGATACGAGCTCACCGCGCGAGACCTCAGTTAGTTGTTGGGGCT 3360  
QY 3349 TACCTAAACACACAGGGTTGCCGCTGCGCAGGACCATCTGAGTTCTGGGAGAGCGTC 3408  
DB 3361 TACCTAAACACACAGGGTTGCCGCTGCGCAGGACCATCTGAGTTCTGGGAGAGCGTC 3420  
QY 3409 TTTACGGGCTCACCCACATAGAGCGGCCATTTCTTTGCCAGACTAAGAGGCGAGGAC 3468

DB 3421 TTTACGGGCTCACCCACATAGACGCCATTTCTTTGTCCGACTAAGAGCGAGAGAC 3480  
QY 3469 AACTTCCCTTACCTGTAGCATACAGGCTACGGTGTGGCGAGGGCTCAGGCTCCACT 3528  
DB 3481 AACTTCCCTTACCTGTAGCATACAGGCTACGGTGTGGCGAGGGCTCAGGCTCCACT 3540  
QY 3529 COATCGTGGGACCAATGTGGAAGTGTCTCATACGGTAAAGCTTACGCTGACGGGCA 3588  
DB 3541 COATCGTGGGACCAATGTGGAAGTGTCTCATACGGTAAAGCTTACGCTGACGGGCA 3600  
QY 3589 ACGCCCTTGTCTATAGGTGGAGCCGTTTCAAAACGAGGTTACTACACACACCCCAT 3648  
DB 3601 ACGCCCTTGTCTATAGGTGGAGCCGTTTCAAAACGAGGTTACTACACACACCCCAT 3660  
QY 3649 ACCAAATACATCATGGCATGCTGCGTGAAGTGGTCACTGGAGGTCGTCAAGAGCACTGGGTG 3708  
DB 3661 ACCAAATACATCATGGCATGCTGCGTGAAGTGGTCACTGGAGCACTGGGTG 3720  
QY 3709 CTGTTAGCGGAGTCTTAGCAGCTCTGGCCGGTATTGCTGACACAGGAGCGGTGTC 3768  
DB 3721 CTGTTAGCGGAGTCTTAGCAGCTCTGGCCGGTATTGCTGACACAGGAGCGGTGTC 3780  
QY 3769 ATTGTGGGAGCATCATCTTGTCCGAAAGCCGCGCATCATTTCCCGACAGGGAAGTCTT 3828  
DB 3781 ATTGTGGGAGCATCATCTTGTCCGAAAGCCGCGCATCATTTCCCGACAGGGAAGTCTT 3840  
QY 3829 TACCGGAGTTCGATGAGATGAAGTGGTGGCTCACACTCCCTTACATCGACAGGGA 3888  
DB 3841 TACCGGAGTTCGATGAGATGAAGTGGTGGCTCACACTCCCTTACATCGACAGGGA 3900  
QY 3889 ATGAGCTCGCGAACAATTTCAAACAGAGGAATCGGTTGCTGCAACAGCCACCAAG 3948  
DB 3901 ATGAGCTCGCGAACAATTTCAAACAGAGGAATCGGTTGCTGCAACAGCCACCAAG 3960  
QY 3949 CAAGCGAGGCTGCTCTCCCGTGGTGAATCAAAGTGGGAGCCCTCGAAGCCTTCTGG 4008  
DB 3961 CAAGCGAGGCTGCTCTCCCGTGGTGAATCAAAGTGGGAGCCCTCGAAGCCTTCTGG 4020  
QY 4009 GCGAAGCATATGGAATTTCAATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTG 4068  
DB 4021 GCGAAGCATATGGAATTTCAATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTG 4080  
QY 4069 CTTGGCAACCCCGGATAGCATCATGATGCAATTCACAGCCTCTATCAGCCGCTC 4128  
DB 4081 CTTGGCAACCCCGGATAGCATCATGATGCAATTCACAGCCTCTATCAGCCGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTTTAAATCCTGGGGGATGGTGGCCGCGCCCAACTTGTCT 4188  
DB 4141 ACCACCAACATACCTCTCTTTAAATCCTGGGGGATGGTGGCCGCGCCCAACTTGTCT 4200  
QY 4189 CTTCCAGCGCTGCTTCTCTTTGTTAGGCGCGGATCGCTGGAGCGGCTGTTGGAGC 4248  
DB 4201 CTTCCAGCGCTGCTTCTCTTTGTTAGGCGCGGATCGCTGGAGCGGCTGTTGGAGC 4260  
QY 4249 ATAGGCTTTGGAGAGTGTCTTGGATATTTTGGCAGGTTATGGAGAGGGTGGCAGGC 4308  
DB 4261 ATAGGCTTTGGAGAGTGTCTTGGATATTTTGGCAGGTTATGGAGAGGGTGGCAGGC 4320  
QY 4309 GCGCTCGTGGCTTTAAGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAAC 4368  
DB 4321 GCGCTCGTGGCTTTAAGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAAC 4380  
QY 4369 CTACTCCCTCTATCTCTCCCTGCGCCCTAGTCTGCGGGTCTGTGCGCAGGATA 4428  
DB 4381 CTACTCCCTCTATCTCTCCCTGCGCCCTAGTCTGCGGGTCTGTGCGCAGGATA 4440  
QY 4429 CTGCTCGGCACTGTTGGGCGGAGGGGCTGTGCTGAGTGAACCGGCTGTAGCG 4488  
DB 4441 CTGCTCGGCACTGTTGGGCGGAGGGGCTGTGCTGAGTGAACCGGCTGTAGCG 4500  
QY 4489 TTCGCTTCGCGGGTAAACACGCTCTCCCGCAGCACTATGTGCTGAGAGCGGCTGCA 4548

Db 4501 TTGCTTCGGGGTAACACAGCTCTCCGCCACACTATGTGCCTGAGAGCAAGCTGCA 4560  
 QY 4549 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAGAGGCTTCAC 4608  
 Db 4561 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAGAGGCTTCAC 4620  
 QY 4609 CAGTGGATCAACAGGAGCTGTCCACGCGCATCTCCGGCTCGTGGCTTAAGAGATGTTGG 4668  
 Db 4621 CAGTGGATCAACAGGAGCTGTCCACGCGCATCTCCGGCTCGTGGCTTAAGAGATGTTGG 4680  
 QY 4669 GATTGGATATGACCGGTGTGATGATTTCAAGACTGTGCTCCAGTCCAAAGTCTCTGCG 4728  
 Db 4681 GATTGGATATGACCGGTGTGATGATTTCAAGACTGTGCTCCAGTCCAAAGTCTCTGCG 4740  
 QY 4729 CGATTGGCGGAGTCCCTCTTCTCATGTCTCAAGTGGGTACAAAGGAGTCTGGCGGGC 4788  
 Db 4741 CGATTGGCGGAGTCCCTCTTCTCATGTCTCAAGTGGGTACAAAGGAGTCTGGCGGGC 4800  
 QY 4789 GACGSCATCATGCAAAACCACTGCCCATGTGGAGCACAGATCACCGGACATGTGAAAAAC 4848  
 Db 4801 GACGSCATCATGCAAAACCACTGCCCATGTGGAGCACAGATCACCGGACATGTGAAAAAC 4860  
 QY 4849 GGTTCATGAGATCTGGGGCTTAGGACCTGTAGTATACAGTGTGCATGGAACATTTCCC 4908  
 Db 4861 GGTTCATGAGATCTGGGGCTTAGGACCTGTAGTATACAGTGTGCATGGAACATTTCCC 4920  
 QY 4909 ATTAAACGGGTACACCAAGCGGCTGTGACGCGCTCCCGGGCCCAAAATTATCTAGGGG 4968  
 Db 4921 ATTAAACGGGTACACCAAGCGGCTGTGACGCGCTCCCGGGCCCAAAATTATCTAGGGG 4980  
 QY 4969 CTGTGGCGGGTGGCTGTGAGGAGTACGTGGAGGTTACGCGGGTGGGGATTTCCACTAC 5028  
 Db 4981 CTGTGGCGGGTGGCTGTGAGGAGTACGTGGAGGTTACGCGGGTGGGGATTTCCACTAC 5040  
 QY 5029 GTGACGGGCATGACCACTGACAAAGTAAAGTCCCGTGTGAGTTCCCGCCCCCGAAATTC 5088  
 Db 5041 GTGACGGGCATGACCACTGACAAAGTAAAGTCCCGTGTGAGTTCCCGCCCCCGAAATTC 5100  
 QY 5089 TTCAAGAGTGGATGGGGTCCGGTTGACAGAGTACGCTCCAGCGTGCAAAACCCCTCTTA 5148  
 Db 5101 TTCAAGAGTGGATGGGGTCCGGTTGACAGAGTACGCTCCAGCGTGCAAAACCCCTCTTA 5160  
 QY 5149 CGGAGGAGGTCAATTCCTGGTCCGGCTCAATCAATACTGGTTGGGTACAGCTCCA 5208  
 Db 5161 CGGAGGAGGTCAATTCCTGGTCCGGCTCAATCAATACTGGTTGGGTACAGCTCCA 5220  
 QY 5209 TGCAGCCGCAACCGGACGTAGAGTGTCTCTTCATGCTCAACGACCCCTCCCAAT 5268  
 Db 5221 TGCAGCCGCAACCGGACGTAGAGTGTCTCTTCATGCTCAACGACCCCTCCCAAT 5280  
 QY 5269 ACGCCGAGACGGCTAAGCGTAGGCTGGCCAGGGATCTCCGCCCTCTTGGCCAGCTCA 5328  
 Db 5281 ACGCCGAGACGGCTAAGCGTAGGCTGGCCAGGGATCTCCGCCCTCTTGGCCAGCTCA 5340  
 QY 5329 TCAGCTAGCAGCTGTCTGGCCCTTCCTTGAAGCAACATGCACTACCCGTCATGACTCC 5388  
 Db 5341 TCAGCTAGCAGCTGTCTGGCCCTTCCTTGAAGCAACATGCACTACCCGTCATGACTCC 5400  
 QY 5389 CCGGACGCTGACCTCATCGAGGCCAACTCTCTGTGGGCGAGGATGGGCGGAAATTC 5448  
 Db 5401 CCGGACGCTGACCTCATCGAGGCCAACTCTCTGTGGGCGAGGATGGGCGGAAATTC 5460  
 QY 5449 ACCCGCTGGAGTCAGAAATAAGGTAGTAAATTTGGACTCTTTCGAGCGGCTCCAGCG 5508  
 Db 5461 ACCCGCTGGAGTCAGAAATAAGGTAGTAAATTTGGACTCTTTCGAGCGGCTCCAGCG 5520  
 QY 5509 GAGGAGATGAGAGGAGTATCCGTTCCGCGGAGATCTCTGGGAGTCCAGGAAATTC 5568  
 Db 5521 GAGGAGATGAGAGGAGTATCCGTTCCGCGGAGATCTCTGGGAGTCCAGGAAATTC 5580  
 QY 5569 CCTGAGCGATGCCATATGGGACCGCGGATTAACCCCTCACTGTTAGATCTCTGG 5628  
 Db 5581 CCTGAGCGATGCCATATGGGACCGCGGATTAACCCCTCACTGTTAGATCTCTGG 5640

QY 5629 AAGGACCCG3ACTAGTCCCTCAGTGGTACACGGGTGTCTCATTTGCCGCTGCCAAGGCC 5688  
 Db 5641 AAGGACCCG3ACTAGTCCCTCAGTGGTACACGGGTGTCTCATTTGCCGCTGCCAAGGCC 5700  
 QY 5689 CCTCGATACACACTCCACGAGGAGAGGAGGTTGCTCTGTGAGATCTACCGTGTCT 5748  
 Db 5701 CCTCGATACACACTCCACGAGGAGAGGAGGTTGCTCTGTGAGATCTACCGTGTCT 5760  
 QY 5749 TCTGCTTTGGGAGTCTCGCCACAAAGACCTTTCCGAGCTCCGAATCTCGGCCCTCGAC 5808  
 Db 5761 TCTGCTTTGGGAGTCTCGCCACAAAGACCTTTCCGAGCTCCGAATCTCGGCCCTCGAC 5820  
 QY 5809 AGCGGACCGCAACGCGCTCTCTGACACAGCCCTCCGACGAGCGGAGTCCGAC 5868  
 Db 5821 AGCGGACCGCAACGCGCTCTCTGACACAGCCCTCCGACGAGCGGAGTCCGAC 5880  
 QY 5869 GTTGAGTCTGTAATCTCCATGCCCTTCCAGGGGAGCGGGGATCCCGATCTCAGC 5928  
 Db 5881 GTTGAGTCTGTAATCTCCATGCCCTTCCAGGGGAGCGGGGATCCCGATCTCAGC 5940  
 QY 5929 GACGGTCTTGGTCTTACCGTAAAGGAGGCTAGTAGAGGCTGTCTGTCTGCTCGATG 5988  
 Db 5941 GACGGTCTTGGTCTTACCGTAAAGGAGGCTAGTAGAGGCTGTCTGTCTGCTCGATG 6000  
 QY 5989 TCCTACATGACGAGCGGCTGATCAGCCATGCGCTGCGGAGGAAACCAAGTGCCT 6048  
 Db 6001 TCCTACATGACGAGCGGCTGATCAGCCATGCGCTGCGGAGGAAACCAAGTGCCT 6060  
 QY 6049 ATCAATGACATGACCACTCTTTGCTCCGTCACCAACTTTGGTCTATGTACAACTCT 6108  
 Db 6061 ATCAATGACATGACCACTCTTTGCTCCGTCACCAACTTTGGTCTATGTACAACTCT 6120  
 QY 6109 CGCAGCGCAAGCTTCCGCGAGAAAGTCACTTTTGACAGACTCGAGTCTTGACGAC 6168  
 Db 6121 CGCAGCGCAAGCTTCCGCGAGAAAGTCACTTTTGACAGACTCGAGTCTTGACGAC 6180  
 QY 6169 CACTACCGGAGTCTCAAGGAGATGAAGCGAGGCTCCAGTTAAGGCTAAATTT 6228  
 Db 6181 CACTACCGGAGTCTCAAGGAGATGAAGCGAGGCTCCAGTTAAGGCTAAATTT 6240  
 QY 6229 CTATCCGTGAGGAGCTGTAACTGACGCGCCCACTTCGGCCAGATCTAAATTTGCG 6288  
 Db 6241 CTATCCGTGAGGAGCTGTAACTGACGCGCCCACTTCGGCCAGATCTAAATTTGCG 6300  
 QY 6289 TATCGGCAAGGAGCTCCGGAACCTATCCAGCAAGGCGGTTAACCACTCCGCTCCGTG 6348  
 Db 6301 TATCGGCAAGGAGCTCCGGAACCTATCCAGCAAGGCGGTTAACCACTCCGCTCCGTG 6360  
 QY 6349 TGGAGGACTTGTGGAAGCACTGAGACACCAATTTGACACCACTCATCGGCAAAAT 6408  
 Db 6361 TGGAGGACTTGTGGAAGCACTGAGACACCAATTTGACACCACTCATCGGCAAAAT 6420  
 QY 6409 GAGTTTTCTCGCTCCAAACAGAGAGGGGGCGCAAGCGAGCTCGCTTATCGTATTC 6468  
 Db 6421 GAGTTTTCTCGCTCCAAACAGAGAGGGGGCGCAAGCGAGCTCGCTTATCGTATTC 6480  
 QY 6469 CCAGATTTGGGGTTCGTGTGCGAGAAATGAGGCTTTTACGATGTTGCTCCACCTC 6528  
 Db 6481 CCAGATTTGGGGTTCGTGTGCGAGAAATGAGGCTTTTACGATGTTGCTCCACCTC 6540  
 QY 6529 CCTCAGGCTGTGAGGCTCTTTCATACGATTTCCATACCTCTCTCGACAGCGGCTGAG 6588  
 Db 6541 CCTCAGGCTGTGAGGCTCTTTCATACGATTTCCATACCTCTCTCGACAGCGGCTGAG 6600  
 QY 6589 TTCTGTGTAATGCTGGAAAGGAGAAATGCGCTATGGGCTTCGCAATGACACCGCG 6648  
 Db 6601 TTCTGTGTAATGCTGGAAAGGAGAAATGCGCTATGGGCTTCGCAATGACACCGCG 6660  
 QY 6649 TGTGTTGACTCAACGCTGACTGAGATGACATCGGCTTTGAGGAGTCAATCTACCAATGT 6708  
 Db 6661 TGTGTTGACTCAACGCTGACTGAGATGACATCGGCTTTGAGGAGTCAATCTACCAATGT 6720

6709 TGTGACTTGGCCCGGAGCCAGACAGAGCCATAAGTGCCTCACAGAGCGGCTTTACATC 6768  
|||||  
6721 TGTGACTTGGCCCGGAGCCAGACAGAGCCATAAGTGCCTCACAGAGCGGCTTTACATC 6780  
|||||  
6769 GGGGGCCCTGACTAATTTCTAAAGGCGAAGACTGCGGCTATGCGCGGTGCGCGCGAGC 6828  
|||||  
6781 GGGGGCCCTGACTAATTTCTAAAGGCGAAGACTGCGGCTATGCGCGGTGCGCGCGAGC 6840  
|||||  
6829 GGTGACTGACACAGCTGCGGTAAATACCTCAATGTTATTGAAGCGCGCTGCGGCC 6888  
|||||  
6841 GGTGACTGACACAGCTGCGGTAAATACCTCAATGTTATTGAAGCGCGCTGCGGCC 6900  
|||||  
6889 TGTGAGCTGCGAAGCTCAGAGCTGCAAGTGTCTGATGCGGACAGCACTTGTCTTT 6948  
|||||  
6901 TGTGAGCTGCGAAGCTCAGAGCTGCAAGTGTCTGATGCGGACAGCACTTGTCTTT 6960  
|||||  
6949 ATCTGTAAAGCGCGGGACCCAAAGAGGACGAGGCTACGGGCTTTCAGGAGGCT 7008  
|||||  
6961 ATCTGTAAAGCGCGGGACCCAAAGAGGACGAGGCTACGGGCTTTCAGGAGGCT 7020  
|||||  
7009 ATGACTAGATCTCTGCCCCCTCGGGACCGGCCCAACACCAAGAAATACGACTTGGAGTTG 7068  
|||||  
7021 ATGACTAGATCTCTGCCCCCTCGGGACCGGCCCAACACCAAGAAATACGACTTGGAGTTG 7080  
|||||  
7069 ATAAATCATGCTCTCCATGTGTCACTGCGGACGATGCTATGCGGAAAGGTTGATC 7128  
|||||  
7081 ATAAATCATGCTCTCCATGTGTCACTGCGGACGATGCTATGCGGAAAGGTTGATC 7140  
|||||  
7129 TATCTCACCGGTGACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC 7188  
|||||  
7141 TATCTCACCGGTGACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC 7200  
|||||  
7189 ACTCCAGTCAATCTCTGGCTAGGCAACATCATCATATGCGCCACCTTGTGGGCAAGG 7248  
|||||  
7201 ACTCCAGTCAATCTCTGGCTAGGCAACATCATCATATGCGCCACCTTGTGGGCAAGG 7260  
|||||  
7249 ATGATCTGATGACTCATTTCTCTCATCTCTAGCTCAGGAACTTGAAGAGCC 7308  
|||||  
7261 ATGATCTGATGACTCATTTCTCTCATCTCTAGCTCAGGAACTTGAAGAGCC 7320  
|||||  
7309 CTAGATTGTGATGCTACGGGGCTGTACTCCATTGAGCCACTTACCTACCTCAGATC 7368  
|||||  
7321 CTAGATTGTGATGCTACGGGGCTGTACTCCATTGAGCCACTTACCTACCTCAGATC 7380  
|||||  
7369 ATTCAACGACTCCATGCGCTTAGCGATTTTACTCCATGTTACTTCCAGGTGATC 7428  
|||||  
7381 ATTCAACGACTCCATGCGCTTAGCGATTTTACTCCATGTTACTTCCAGGTGATC 7440  
|||||  
7429 AATAGGCTGCTTCATGCGCTCAGGAACTTGGGTTACCGCCCTTGGAGTCTGGAGCAT 7488  
|||||  
7441 AATAGGCTGCTTCATGCGCTCAGGAACTTGGGTTACCGCCCTTGGAGTCTGGAGCAT 7500  
|||||  
7489 CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGC 7548  
|||||  
7501 CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGC 7560  
|||||  
7549 AAGTACTCTTCAACTGGGAGTAAGACCAAGCTCAAACTCACTCOAATCCCGGCTGGG 7608  
|||||  
7561 AAGTACTCTTCAACTGGGAGTAAGACCAAGCTCAAACTCACTCOAATCCCGGCTGGG 7620  
|||||  
7609 TCCAGTTGATTTATCCAGCTGGTGTGTTGTTGATGCGGGGAGACATATATCAC 7668  
|||||  
7621 TCCAGTTGATTTATCCAGCTGGTGTGTTGTTGATGCGGGGAGACATATATCAC 7680  
|||||  
7669 AGCTGTCTGTGCGCGACCCGCTGTTCATGTGGTGCCTACTCTCTACTTTCTGTAGGG 7728  
|||||  
7681 AGCTGTCTGTGCGCGACCCGCTGTTCATGTGGTGCCTACTCTCTACTTTCTGTAGGG 7740  
|||||  
7729 GTAGGCTATCTACTATCTCCCAACCGGATGAACCGGGACCTAAACACTCCAGGCCAATAGG 7788  
|||||  
7741 GTAGGCTATCTACTATCTCCCAACCGGATGAACCGGGAGCTAAACACTCCAGGCCAATAGG 7800  
|||||  
7789 CCATCTGTTTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7848

Db 7801 CCATCTGTTTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
|||||  
QY 7849 TTTTCTCTTTTCTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTT 7908  
|||||  
Db 7861 TTTTCTCTTTTCTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTTCTCTTTT 7920  
|||||  
QY 7909 CCTAGTCAAGCTAGCTGTGAAGCTCGTGAAGCTCGTGAAGCTCGTGAAGCTCGTGAAGCTCGTGAAGCT 7968  
|||||  
Db 7921 CCTAGTCAAGCTAGCTGTGAAGCTCGTGAAGCTCGTGAAGCTCGTGAAGCTCGTGAAGCTCGTGAAGCT 7980  
|||||  
QY 7969 TGGCCTCTCTGAGATCAAGT 7989  
|||||  
Db 7981 TGGCCTCTCTGAGATCAAGT 8001  
|||||  
RESULT 7  
SSE242654 8001 bp RNA linear SYN 11-MAY-2000  
LOCUS Hepatitis C virus replicon I389/NS3-3'UTR.  
DEFINITION AJ242654  
ACCESSION AJ242654.1 GI:5441840  
VERSION core-neo gene; NS2 gene; NS2 proteinase; NS3 gene; NS3  
KEYWORDS proteinase/helicase; NS3/4A proteinase cofactor; NS4A gene; NS4B  
gene; NS4B protein; NS5a gene; NS5A phosphoprotein; NS5b gene; NS5B  
RNA dependant RNA polymerase; polyprotein.  
SOURCE Hepatitis C virus replicon I389/NS3-3'UTR  
ORGANISM Hepatitis C virus replicon I389/NS3-3'UTR  
REFERENCE artificial sequences; vectors.  
1  
AUTHORS Lohmann,V., Korner,F., Koch,J., Herian,U., Theilmann,L. and  
Bartenschlager,R.  
TITLE Replication of subgenomic hepatitis C virus RNAs in a hepatoma cell  
line  
JOURNAL Science 285 (5424), 110-113 (1999)  
MEDLINE 99322193  
PUBMED 10390360  
REFERENCE 2 (bases 1 to 8001)  
AUTHORS Bartenschlager,R.  
TITLE Direct Submission  
JOURNAL Submitted (26-MAY-1999) Bartenschlager R., Institute for Virology,  
Johannes Gutenberg - University Mainz, Obere Zahlbacher Strasse 67,  
55131 Mainz, GERMANY  
FEATURES  
Location/Qualifiers  
1..8001  
/organism="Hepatitis C virus replicon I389/NS3-3'UTR"  
/mol\_type="other RNA"  
/db\_xref="taxon:95365"  
/focus  
/note="Neomycin-selectable bicistronic subgenomic  
hepatitis C virus replicon"  
join(1..341,1813..7770)  
/organism="Hepatitis C virus"  
/mol\_type="other RNA"  
/strain="1b"  
/isolate="Con1"  
/db\_xref="taxon:11103"  
/tissue\_type="liver"  
/country="Germany"  
1202..1812  
/organism="Encephalomyocarditis virus"  
/mol\_type="other RNA"  
/db\_xref="taxon:12104"  
1..341  
/evidence=experimental  
342..1193  
/gene="core-neo"  
342..1193  
/genes="core-neo"  
/function="selectable marker"  
/codon\_start=1  
/product="core-neo fusion protein"  
/protein\_id="CAB46916.1"





Db 961 ATGGCGATGCCGTCTGGCGAATATCATGGTGGAAATAGGCCGCTTTTCTGGATTCATCG 1020  
QY  
1009 ACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATA 1068  
Db  
1021 ACTGTGGCCGGCTGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATA 1080  
QY  
1069 TTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTGTGCTTTACGGTATCGCGG 1128  
Db  
1081 TTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCTGTGCTTTACGGTATCGCGG 1140  
QY  
1129 CTCGCCATTCGACGCGCATCGCCTTCTATCGCCTTCTTCAACAGTTCCTTCTGAGTTAAA 1188  
Db  
1141 CTCGCCATTCGACGCGCATCGCCTTCTATCGCCTTCTTCAACAGTTCCTTCTGAGTTAAA 1200  
QY  
1189 CAGACCAACAGGTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCT 1248  
Db  
1201 CAGACCAACAGGTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCT 1260  
QY  
1249 AAGCTTACTGGCGGAAGCCGCTTGGAAATAGGCGCGGTGTGGTTGTCTATATGTTATTT 1308  
Db  
1261 AAGCTTACTGGCGGAAGCCGCTTGGAAATAGGCGCGGTGTGGTTGTCTATATGTTATTT 1320  
QY  
1309 TCCACCATATGCCGTCTTTTGGCAATGAGGCGCCGAAACCTGGGCCCTGTCTCTTTG 1368  
Db  
1321 TCCACCATATGCCGTCTTTTGGCAATGAGGCGCCGAAACCTGGGCCCTGTCTCTTTG 1380  
QY  
1369 AGGAGATTCCTAGGGGTCTTTCCCTCTCGCAAGAAATGCAAGTCTGTGTGAATGTC 1428  
Db  
1381 ACGAGATTCCTAGGGGTCTTTCCCTCTCGCAAGAAATGCAAGTCTGTGTGAATGTC 1440  
QY  
1429 GTGAAGAAAGCTTCTCTGGAAGCTTCTTGAAGCAAAACAAAGTCTGTAGGACCCCTT 1488  
Db  
1441 GTGAAGAAAGCTTCTCTGGAAGCTTCTTGAAGCAAAACAAAGTCTGTAGGACCCCTT 1500  
QY  
1489 TCGAGCAGCGGAACCCCTCTCGGACAGTGCTCTCGGCCCAAGCCACAGCTGA 1548  
Db  
1501 TCGAGCAGCGGAACCCCTCTCGGACAGTGCTCTCGGCCCAAGCCACAGCTGA 1560  
QY  
1549 TAAGATACACCTGMAAGGCGGCACAACCCAGTGCACAGTTGTGAGTTGGATGTTGTG 1608  
Db  
1561 TAAGATACACCTGMAAGGCGGCACAACCCAGTGCACAGTTGTGAGTTGGATGTTGTG 1620  
QY  
1609 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAAGGGGTGAAGGATGCCAGAAG 1668  
Db  
1621 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAAGGGGTGAAGGATGCCAGAAG 1680  
QY  
1669 GTACCCCATTTGATGGATCTGATCTGGGGCTCTCGGTGCACATGCTTTACATGTTTAG 1728  
Db  
1681 GTACCCCATTTGATGGATCTGATCTGGGGCTCTCGGTGCACATGCTTTACATGTTTAG 1740  
QY  
1729 TCGAGTTAAAAACGCTTAGGCCCCCGAACACCGGGGACGTTGTTTCTTTGAAAAA 1788  
Db  
1741 TCGAGTTAAAAACGCTTAGGCCCCCGAACACCGGGGACGTTGTTTCTTTGAAAAA 1800  
QY  
1789 CAGGATAATACCATGGCGCTATTAGGCGTACTCCACAGACGCGAGCCCTACTTGGC 1848  
Db  
1801 CAGGATAATACCATGGCGCTATTAGGCGTACTCCACAGACGCGAGCCCTACTTGGC 1860  
QY  
1849 TGCATCATCACTAGCTCTCAGSCCGGACAGGAACAGTCCAGGGGAGGTCCCAAGTG 1908  
Db  
1861 TGCATCATCACTAGCTCTCAGSCCGGACAGGAACAGTCCAGGGGAGGTCCCAAGTG 1920  
QY  
1909 GTCTCCACCGCAACAATCTTTCTCGGACCTGGCTCAATGGCGTGTGTGGACTGTC 1968  
Db  
1921 GTCTCCACCGCAACAATCTTTCTCGGACCTGGCTCAATGGCGTGTGTGGACTGTC 1980  
QY  
1969 TATCATGGTCCGGCTCAAGAACCTTTGCGGCCCAAGGGCCCAATCAACCAAAATGTAC 2028  
Db  
1981 TATCATGGTCCGGCTCAAGAACCTTTGCGGCCCAAGGGCCCAATCAACCAAAATGTAC 2040  
QY  
2029 ACCAATGTGACAGGACCTCGTCTGGTGGCAAGCGCCCGCGGCGCTTCTCTTGACA 2088

Db 2041 ACCAATGTGACAGGACCTCGTCTGGTGGCAAGCGCCCGCGGCGGCTTCTCTTGACA 2100  
QY  
2089 COATSCA CTTGCGGAGCTCGGACCTTTACTTTGTGTCA CAGAGCATGCGGATGTCATTCG 2148  
Db  
2101 CCATGCA CTTGCGGAGCTCGGACCTTTACTTTGTGTCA CAGAGCATGCGGATGTCATTCG 2160  
QY  
2149 GTGCGCGCGGCGGCGAGCAGCAGGGGAGCTTACTCTCCCCAGGCGCGCTCTCTACTTTG 2208  
Db  
2161 GTGCGCGCGGCGGCGAGCAGCAGGGGAGCTTACTCTCCCCAGGCGCGCTCTCTACTTTG 2220  
QY  
2209 AAGGGCTCTTCTGCGGCGGCTCCACTGTCTGCCCCCTCGGGSCAGCTGTGGGCATCTTTG 2268  
Db  
2221 AAGGGCTCTTCTGCGGCGGCTCCACTGTCTGCCCCCTCGGGSCAGCTGTGGGCATCTTTG 2280  
QY  
2269 GTGCGCGGTGTGACCCGAGGGGTTCGGAAGGGGTGGACTTTGTACCCGTGAGTCTATG 2328  
Db  
2281 GCTGCGGTGTGACCCGAGGGGTTCGGAAGGGGTGGACTTTGTACCCGTGAGTCTATG 2340  
QY  
2329 GAAACCACTATGCGGTCCCGGTCTTTCAGGCAAACTCGTCCCTCCCGCGCTACCGCAG 2388  
Db  
2341 GAAACCACTATGCGGTCCCGGTCTTTCAGGCAAACTCGTCCCTCCCGCGCTACCGCAG 2400  
QY  
2389 ACAATCCAGGTGGCCCATCTACACGCCCTACTGTGTAGCGGCAAGACATAAGGTGCGG 2448  
Db  
2401 ACAATCCAGGTGGCCCATCTACACGCCCTACTGTGTAGCGGCAAGACATAAGGTGCGG 2460  
QY  
2449 GTGCGGTATGCAAGCCCAAGGGTATTAAGGTGCTTTGCTGTGAACCCGTCTCGCGCCAC 2508  
Db  
2461 GCTGCGTATGCAAGCCCAAGGGTATTAAGGTGCTTTGCTGTGAACCCGTCTCGCGCCAC 2520  
QY  
2509 CTAGGTTTCGGGCGCTATATGCTTAAGGCACATGCTATCGACCCCTAAACATCAGAACCGG 2568  
Db  
2521 CTAGGTTTCGGGCGCTATATGCTTAAGGCACATGCTATCGACCCCTAAACATCAGAACCGG 2580  
QY  
2569 GTAAGGACCATCACACCGGTGCCCTCATCAGCTACTCCACCTATGGCAAGTTCTTTGCC 2628  
Db  
2581 GTAAGGACCATCACACCGGTGCCCTCATCAGCTACTCCACCTATGGCAAGTTCTTTGCC 2640  
QY  
2629 GACGTTGTTGCTCTGGGGCGCTATATGATCATATAATATGATGAGTGCACTCAACT 2688  
Db  
2641 GACGTTGTTGCTCTGGGGCGCTATATGATCATATAATATGATGAGTGCACTCAACT 2700  
QY  
2689 GACTCGACACTATCTGGGCGCTATGCTTAAGGCACATGCTATCGACCCCTAAACATCAGAACCGG 2748  
Db  
2701 GACTCGACACTATCTGGGCGCTATGCTTAAGGCACATGCTATCGACCCCTAAACATCAGAACCGG 2760  
QY  
2749 CGACTCGTCTGTGCCACCGCTACGCTCCGGGATCGGTCAACCGTGCCACATCCAAAC 2808  
Db  
2761 CGACTCGTCTGTGCCACCGCTACGCTCCGGGATCGGTCAACCGTGCCACATCCAAAC 2820  
QY  
2809 ATCGAGGAGTGGCTCTGTGCGACTGGAGAAAATCCCTTTTATGGGAAAGCATCCCT 2868  
Db  
2821 ATCGAGGAGTGGCTCTGTGCGACTGGAGAAAATCCCTTTTATGGGAAAGCATCCCT 2880  
QY  
2869 ATCGAGACCATCAAGGGGCGGAGGACCTCATTTCTGCCATTTCCAAAGAAAGTCTGAT 2928  
Db  
2881 ATCGAGACCATCAAGGGGCGGAGGACCTCATTTCTGCCATTTCCAAAGAAAGTCTGAT 2940  
QY  
2929 GAGCTCGCGCGAAGCTGTCCGCGCTCGGACTCAATGCTGTAGCATATTTACCGGGGCTT 2988  
Db  
2941 GAGCTCGCGCGAAGCTGTCCGCGCTCGGACTCAATGCTGTAGCATATTTACCGGGGCTT 3000  
QY  
2989 GATGTATCCGTATACCAACTAGCGGAGAGCTCATTTGTCTGTAGCAACCGAGCTCTTAATG 3048  
Db  
3001 GATGTATCCGTATACCAACTAGCGGAGAGCTCATTTGTCTGTAGCAACCGAGCTCTTAATG 3060  
QY  
3049 ACGGGCTTTACCGGCGATTTTCGACTCAGTGATCGACTGCAATACATGTTCTCCACAGACA 3108  
Db  
3061 ACGGGCTTTACCGGCGATTTTCGACTCAGTGATCGACTGCAATACATGTTCTCCACAGACA 3120  
QY  
3109 GTGACTTTCAGCTGAGCCCGACCTTTCACATTTGAGACGACGACCGCTGCGCAAGACGCG 3168  
Db  
3121 GTGACTTTCAGCTGAGCCCGACCTTTCACATTTGAGACGACGACCGCTGCGCAAGACGCG 3180

QY	3169	GTGTCAACGCTCGACGGCGAGGACAGTGGTAGGGCAGAGATGGGCATTTACAGGTTT	3228
Db	3181	GTGTCAACGCTCGACGGCGAGGACAGTGGTAGGGCAGAGATGGGCATTTACAGGTTT	3240
QY	3229	GTGACTCCAGGAGAACGGCCCTCGGGCATTTGCGATTCTCGGTTCTGTGCGAGTGCTAT	3288
Db	3241	GTGACTCCAGGAGAACGGCCCTCGGGCATTTGCGATTCTCGGTTCTGTGCGAGTGCTAT	3300
QY	3289	GACGGGGCTGTGCTTGTGTACAGAGCTCAGCCCGCCGAGACCTCAGTTAGGTTGCGGCT	3348
Db	3301	GACGGGGCTGTGCTTGTGTGTACAGAGCTCAGCCCGCCGAGACCTCAGTTAGGTTGCGGCT	3360
QY	3349	TACCTTAAACACACACAGGTTGCGCTCTGCGAGGACCATCTGGAGTTCTGGAGAGCGTC	3408
Db	3361	TACCTTAAACACACACAGGTTGCGCTCTGCGAGGACCATCTGGAGTTCTGGAGAGCGTC	3420
QY	3409	TTTACAGCCCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGACAGAGAC	3468
Db	3421	TTTACAGCCCTCACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGACAGAGAC	3480
QY	3469	AACCTCCCTACCTGGTAGCATACAGAGTAGGTTGCGGCGAGGGCTCAGGCTCCACT	3528
Db	3481	AACCTCCCTACCTGGTAGCATACAGAGTAGGTTGCGGCGAGGGCTCAGGCTCCACT	3540
QY	3529	CCATCGTGGGACCAAACTGTGAAGTGTCTCATAGCGCTAAAGCCTACGCTGCACGGGCA	3588
Db	3541	CCATCGTGGGACCAAACTGTGAAGTGTCTCATAGCGCTAAAGCCTACGCTGCACGGGCA	3600
QY	3589	ACGCCCTCTGTGTATPAGGCTGGAGCGTTCAAAACGAGGTTACTACACACACCCCCATA	3648
Db	3601	ACGCCCTCTGTGTATPAGGCTGGAGCGTTCAAAACGAGGTTACTACACACACCCCCATA	3660
QY	3649	ACCAATATACATATGCGATGCTGCGCTGACCTGGAGGTGCTCACAGACACTGGGTG	3708
Db	3661	ACCAATATACATATGCGATGCTGCGCTGACCTGGAGGTGCTCACAGACACTGGGTG	3720
QY	3709	CTGCTPAGGCGGAGTCTTAGCAGCTCTGGCGCGTATTTCCTGACAAACAGGAGGTTGTC	3768
Db	3721	CTGCTPAGGCGGAGTCTTAGCAGCTCTGGCGCGTATTTCCTGACAAACAGGAGGTTGTC	3780
QY	3769	ATTGTGGGAGGATCATCTTGTCCGGAAGCGGCCATCATTTCCCGACAGGGAAGTCTTT	3828
Db	3781	ATTGTGGGAGGATCATCTTGTCCGGAAGCGGCCATCATTTCCCGACAGGGAAGTCTTT	3840
QY	3829	TACGGGAGTTTCGATGATGAGAGTGCSCCTCACACCTCCCTTACATCGAACAGGGA	3888
Db	3841	TACGGGAGTTTCGATGATGAGAGTGCSCCTCACACCTCCCTTACATCGAACAGGGA	3900
QY	3889	ATGCAGCTCGCCGAAACAAATTCAAACAGAGGCAATCGGGTTGCTGCAAAAGCCACCAAG	3948
Db	3901	ATGCAGCTCGCCGAAACAAATTCAAACAGAGGCAATCGGGTTGCTGCAAAAGCCACCAAG	3960
QY	3949	CAACGGAGGCTGTGCTCCGTTGGTGAATCCAAAGTGGGAGCCCTCGAAGCCTTCTGG	4008
Db	3961	CAACGGAGGCTGTGCTCCGTTGGTGAATCCAAAGTGGGAGCCCTCGAAGCCTTCTGG	4020
QY	4009	GCGAAGCATATGTGGAATTTTCATCAGCGGGATACAATATTTAGCAGGCTTTCACACTG	4068
Db	4021	GCGAAGCATATGTGGAATTTTCATCAGCGGGATACAATATTTAGCAGGCTTTCACACTG	4080
QY	4069	CCTGGCAACCCCGGATAGCATCACTGATGCAATTCACAGCCTCTATCACCAGCCCGCTC	4128
Db	4081	CCTGGCAACCCCGGATAGCATCACTGATGCAATTCACAGCCTCTATCACCAGCCCGCTC	4140
QY	4129	ACCAACCAATACCTCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGCT	4188
Db	4141	ACCAACCAATACCTCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGCT	4200
QY	4189	CCTCCAGCGCTGTTCTGCTTTCTGTTAGGGCCGGCATCGTGGAGGGCTGTTGGCAGC	4248
Db	4201	CCTCCAGCGCTGTTCTGCTTTCTGTTAGGGCCGGCATCGTGGAGGGCTGTTGGCAGC	4260

QY	4249	ATAGCCCTTGGAAAGGTGCTTTGTGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGC	4308
Db	4261	ATAGCCCTTGGAAAGGTGCTTTGTGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGC	4320
QY	4309	CGCCTCGTGGCTTTTAAGGTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGTTTAC	4368
Db	4321	CGCCTCGTGGCTTTTAAGGTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGTTTAC	4380
QY	4369	CTACTCCCTGTATCTCTCCCTGGCGCCCTTAGTCTCGGGTCTGTGCGACGATA	4428
Db	4381	CTACTCCCTGTATCTCTCCCTGGCGCCCTTAGTCTCGGGTCTGTGCGACGATA	4440
QY	4429	CTGCGTCGCGACGTGGGCCAGGGGAGGGGCTGTGCAAGTGGATGAAACCGCTGTATGG	4488
Db	4441	CTGCGTCGCGACGTGGGCCAGGGGAGGGGCTGTGCAAGTGGATGAAACCGCTGTATGG	4500
QY	4489	TTGCGTTCCGCGGGTAAACACAGTCTCCCGCCAGCACTATGTGCCTGAGAGCGACCTGCA	4548
Db	4501	TTGCGTTCCGCGGGTAAACACAGTCTCCCGCCAGCACTATGTGCCTGAGAGCGACCTGCA	4560
QY	4549	GCAAGTGTCACTCAGATCCTCTAGTCTTACCAATCACTCAGCTGCTGAAAGGCTTCA	4608
Db	4561	GCAAGTGTCACTCAGATCCTCTAGTCTTACCAATCACTCAGCTGCTGAAAGGCTTCA	4620
QY	4609	CAGTGGATCAACGAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTTAAGAGTCTTGG	4668
Db	4621	CAGTGGATCAACGAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTTAAGAGTCTTGG	4680
QY	4669	GATTGGATATACAGCGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCCAAGCTCTCGCG	4728
Db	4681	GATTGGATATACAGCGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCCAAGCTCTCGCG	4740
QY	4729	CGATTGCGGAGTCCCTCTTCTCATGTCAAAGTGGGTACAAGGAGTCTGGCGGGC	4788
Db	4741	CGATTGCGGAGTCCCTCTTCTCATGTCAAAGTGGGTACAAGGAGTCTGGCGGGC	4800
QY	4789	GACGCGCATATGCAAAACACCTGCCCATGTGGAGCACAGATCACGGACATGTGAAAC	4848
Db	4801	GACGCGCATATGCAAAACACCTGCCCATGTGGAGCACAGATCACGGACATGTGAAAC	4860
QY	4849	GGTTCCATGAGGATCGTGGGCTTAGGACCTGTAGTAACACGTGGCATGGAACATTTCCC	4908
Db	4861	GGTTCCATGAGGATCGTGGGCTTAGGACCTGTAGTAACACGTGGCATGGAACATTTCCC	4920
QY	4909	ATTAAACGGGTACACACCGGGCCCTGACGCGCTCCCGCGGCCAAATATTCTAGGCG	4968
Db	4921	ATTAAACGGGTACACACCGGGCCCTGACGCGCTCCCGCGGCCAAATATTCTAGGCG	4980
QY	4969	CTGTGGCGGTGGCTGTGAGGAGTACGTGGAGGTACGCGGGTGGGGATTTCCACTAC	5028
Db	4981	CTGTGGCGGTGGCTGTGAGGAGTACGTGGAGGTACGCGGGTGGGGATTTCCACTAC	5040
QY	5029	GTGACGGGCTATGACCACTGACCAAGTAAAGTGCCTGTGTGAGGTTCCGGCCCCCGAATTC	5088
Db	5041	GTGACGGGCTATGACCACTGACCAAGTAAAGTGCCTGTGTGAGGTTCCGGCCCCCGAATTC	5100
QY	5089	TTTACACAGTGGATGGGTGCGGTTCGACAGGTACGCTCCAGCGTGAACCCCTCTTA	5148
Db	5101	TTTACACAGTGGATGGGTGCGGTTCGACAGGTACGCTCCAGCGTGAACCCCTCTTA	5160
QY	5149	CGGAGGAGGTTCATTTCTTGGTGGGTCAATCAATACCTGTTGGGTCAAGCTCCCA	5208
Db	5161	CGGAGGAGGTTCATTTCTTGGTGGGTCAATCAATACCTGTTGGGTCAAGCTCCCA	5220
QY	5209	TGGAGCCCGAAGCGGACGTAGCAGTGTCTTCCATGTCTACCGACCCCTCCCACTT	5268
Db	5221	TGGAGCCCGAAGCGGACGTAGCAGTGTCTTCCATGTCTACCGACCCCTCCCACTT	5280
QY	5269	ACGGCGAGAGCGGTAGCGTGGCCAGGGGATCTCCCGCTCTCTGGCGAGCTCA	5328
Db	5281	ACGGCGAGAGCGGTAGCGTGGCCAGGGGATCTCCCGCTCTCTGGCGAGCTCA	5340
QY	5329	TCAGTAGCCAGGTGTCTGGCGCTTCTTGTGAAAGCAACATGACACTACCGCTCATGCTCC	5388

Db 5341 TCAGTAGCCAGCTCTCTGGGCTTCTTTGAAGGCAACATGCACATACCGCTCATGACTCC 5400  
QY 5389 CCGGACGCTGACCTCATCGAGGCCAACCTCTCTGGCGGAGAGATGGCGGGAACATC 5448  
Db 5401 CCGGACGCTGACCTCATCGAGGCCAACCTCTCTGGCGGAGAGATGGCGGGAACATC 5460  
QY 5449 ACCCGGTGGAGTCAGAAATAAGGTAGTAATTTTGGACTCTTTTGAGCCGCTCCAAAGG 5508  
Db 5461 ACCCGGTGGAGTCAGAAATAAGGTAGTAATTTTGGACTCTTTTGAGCCGCTCCAAAGG 5520  
QY 5509 GAGGAGGATGAGAGGAGTAGTCCGTTCCGGCGGAGATCCTCGGAGGTCAGGAAATTC 5568  
Db 5521 GAGGAGGATGAGAGGAGTAGTCCGTTCCGGCGGAGATCCTCGGAGGTCAGGAAATTC 5580  
QY 5569 CCTCAGCGATGCCCATATGGSCACGCCCGGATTAACAACCTCCACTGTAGAGTCCTGG 5628  
Db 5581 CCTCAGCGATGCCCATATGGSCACGCCCGGATTAACAACCTCCACTGTAGAGTCCTGG 5640  
QY 5629 AAGGACCGGACTACGTCCTCCAGTGTGATACAGGGTGTCCATTCGCGCTGCCAAGGCC 5688  
Db 5641 AAGGACCGGACTACGTCCTCCAGTGTGATACAGGGTGTCCATTCGCGCTGCCAAGGCC 5700  
QY 5689 CCTCGATACCACTCCAGGAGGAGGACGTTGTCTCTCAGAACTACCGTGTCT 5748  
Db 5701 CCTCGATACCACTCCAGGAGGAGGACGTTGTCTCTCAGAACTACCGTGTCT 5760  
QY 5749 TCTGCTTTGGCGAGCTGCCACAAAGACCTTCGCGAGCTCCGAATCTGTCGCGCTCGAC 5808  
Db 5761 TCTGCTTTGGCGAGCTGCCACAAAGACCTTCGCGAGCTCCGAATCTGTCGCGCTCGAC 5820  
QY 5809 AGCGGACGGCAAGCGCTCTCTGACAGCCCTCGAGCAGCGGACCGGGATCCGAC 5868  
Db 5821 AGCGGACGGCAAGCGCTCTCTGACAGCCCTCGAGCAGCGGACCGGGATCCGAC 5880  
QY 5869 GTTAGTCTGACTCTCTCCATGCCCCCTTTGAGGGGAGCGGGGATCCCGATCTCAGC 5928  
Db 5881 GTTAGTCTGACTCTCTCCATGCCCCCTTTGAGGGGAGCGGGGATCCCGATCTCAGC 5940  
QY 5929 GACGGTCTTGTCTACCGTAAGCAGGAGGCTAGTAGAGAGCTGCTCTGCTCGATG 5988  
Db 5941 GACGGTCTTGTCTACCGTAAGCAGGAGGCTAGTAGAGAGCTGCTCTGCTCGATG 6000  
QY 5989 TCCTACATGACAGGCGCCCTGATCAGCCATGCGCTGCGGAGGAACCAAGTGGCC 6048  
Db 6001 TCCTACATGACAGGCGCCCTGATCAGCCATGCGCTGCGGAGGAACCAAGTGGCC 6060  
QY 6049 ATCAATGCACTGAGCAACTCTTTGCTCCGTCACCACTTGGTCTATGCTACACATCT 6108  
Db 6061 ATCAATGCACTGAGCAACTCTTTGCTCCGTCACCACTTGGTCTATGCTACACATCT 6120  
QY 6109 CGCAGCGAAGCTCGGCGAGAGAGGTACCTTTGACAGACTGCAGGTCCTGGACGAC 6168  
Db 6121 CGCAGCGAAGCTCGGCGAGAGAGGTACCTTTGACAGACTGCAGGTCCTGGACGAC 6180  
QY 6169 CACTACCGGAGCTGCTCAAGAGATGAAGGAGGCGTCCACAGTTAAGGTAACTT 6228  
Db 6181 CACTACCGGAGCTGCTCAAGAGATGAAGGAGGCGTCCACAGTTAAGGTAACTT 6240  
QY 6229 CTATCCGTGAGGAGCTGTAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGC 6288  
Db 6241 CTATCCGTGAGGAGCTGTAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGC 6300  
QY 6289 TATGGGCAAGAGCTCGGCAACCTATCAGCAAGCGCGTTAAACCACATCCGCTCCGTG 6348  
Db 6301 TATGGGCAAGAGCTCGGCAACCTATCAGCAAGCGCGTTAAACCACATCCGCTCCGTG 6360  
QY 6349 TGAAGGACTTGTGTAAGACACTGACACCAATTTGACCAACCAATATGGAAGGCAAAAT 6408  
Db 6361 TGAAGGACTTGTGTAAGACACTGACACCAATTTGACCAACCAATATGGAAGGCAAAAT 6420  
QY 6409 GAGGTTTCTGGTCCAAACGAGAGGGGGCGGACGCTGCCTTATCGTATTC 6468

Db 6421 GAGGTTTCTGGTCCAAACAGAGAGGGGGCCCAAGCCAGCTCGCTTATCGTATTC 6480  
QY 6469 CCAGATTTGGGGTTTGTGTGTGCGAGAAAATGGCCCTTTACGATGTGTGTCTCCACCTTC 6528  
Db 6481 CCAGATTTGGGGTTTGTGTGTGCGAGAAAATGGCCCTTTACGATGTGTGTCTCCACCTTC 6540  
QY 6529 CCTCAGGCGTGTGGGCTCTTCATACGATTCGAATCTCCTCTGAGCAGCGGTGCGAG 6588  
Db 6541 CCTCAGGCGTGTGGGCTCTTCATACGATTCGAATCTCCTCTGAGCAGCGGTGCGAG 6600  
QY 6589 TTCTCTGTGAATGCTCTGAAAACGGAATAATGCCCTATGGGCTTCGCATATGACACCCGC 6648  
Db 6601 TTCTCTGTGAATGCTCTGAAAACGGAATAATGCCCTATGGGCTTCGCATATGACACCCGC 6660  
QY 6649 TGTTTTGACTCAACCGGTCACTGAGAAATGACATPCGTGTGTGAGGAGTCAATCTCAAAATG 6708  
Db 6661 TGTTTTGACTCAACCGGTCACTGAGAAATGACATPCGTGTGTGAGGAGTCAATCTCAAAATG 6720  
QY 6709 TGTGACTTTGGCCCCCGAAGCCAGACAGGCTCAATAGGTCTCACAGAGCGGCTTTACATC 6768  
Db 6721 TGTGACTTTGGCCCCCGAAGCCAGACAGGCTCAATAGGTCTCACAGAGCGGCTTTACATC 6780  
QY 6769 GGGGGCCCCCTGACTTAATTTCTAAAGGGCAGAACTCGCGGCTATCGCGGTGCGCGGAGC 6828  
Db 6781 GGGGGCCCCCTGACTTAATTTCTAAAGGGCAGAACTCGCGGCTATCGCGGTGCGCGGAGC 6840  
QY 6829 GGTGTACTGACGACGAGCTGCGGTAAATACCTTCAATGTCTTAAAGGCGCGCTGCGGCC 6888  
Db 6841 GGTGTACTGACGACGAGCTGCGGTAAATACCTTCAATGTCTTAAAGGCGCGCTGCGGCC 6900  
QY 6889 TGTGAGCTGGGAGCTCCAGGACTGCAGATGCTGATGCGGAGACGACCTTGTCTGT 6948  
Db 6901 TGTGAGCTGGGAGCTCCAGGACTGCAGATGCTGATGCGGAGACGACCTTGTCTGT 6960  
QY 6949 ATCTCTGAAAAGCGGGGACCCCAAGAGGACGAGGCGAGCTACGGGCTTCAACGAGGCT 7008  
Db 6961 ATCTCTGAAAAGCGGGGACCCCAAGAGGACGAGGCGAGCTACGGGCTTCAACGAGGCT 7020  
QY 7009 ATGACTAGATCTCTGCCCCCTCTGGGACCCGCGCAACCAAGAAATACGACTTGGAGTTG 7068  
Db 7021 ATGACTAGATCTCTGCCCCCTCTGGGACCCGCGCAACCAAGAAATACGACTTGGAGTTG 7080  
QY 7069 ATAAATCATGCTCTCCAAATGTGTGCTGCGCAGCATGCTTGGGAAAAGGTTGATC 7128  
Db 7081 ATAAATCATGCTCTCCAAATGTGTGCTGCGCAGCATGCTTGGGAAAAGGTTGATC 7140  
QY 7129 TATCTCACCCGTGACCCCAACCAACCCCTTGGCGGGCTGCGTGGGAGACAGTACAC 7188  
Db 7141 TATCTCACCCGTGACCCCAACCAACCCCTTGGCGGGCTGCGTGGGAGACAGTACAC 7200  
QY 7189 ACTCAGTCAATTCCTGGCTAGGCAACATCATATGATGCGGCCACCTTGTGGGCAAG 7248  
Db 7201 ACTCAGTCAATTCCTGGCTAGGCAACATCATATGATGCGGCCACCTTGTGGGCAAG 7260  
QY 7249 ATGATCTGATGACTCATTTCTTCCATCTTCTAGCTCAGGAAACACTTGAAGAGCC 7308  
Db 7261 ATGATCTGATGACTCATTTCTTCCATCTTCTAGCTCAGGAAACACTTGAAGAGCC 7320  
QY 7309 CTAGATTTGTGATCTACGGGCTGTTACTCTCATTTGAGCCACTTGACCTACAGATC 7368  
Db 7321 CTAGATTTGTGATCTACGGGCTGTTACTCTCATTTGAGCCACTTGACCTACAGATC 7380  
QY 7369 ATTCAACGACTCCATGGCTTAGCGGATTTTCACTCCATAGTTACTCTCCAGGTAGATC 7428  
Db 7381 ATTCAACGACTCCATGGCTTAGCGGATTTTCACTCCATAGTTACTCTCCAGGTAGATC 7440  
QY 7429 AATAGGTGGCTTCACTGCTCAGGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACAT 7488  
Db 7441 AATAGGTGGCTTCACTGCTCAGGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACAT 7500  
QY 7489 CGGGCCAGAAAGTGTCCGCGTAGGCTACTGTCCCGAGGGGGGAGGCTGCCACTTGTGGC 7548  
Db 7501 CGGGCCAGAAAGTGTCCGCGTAGGCTACTGTCCCGAGGGGGGAGGCTGCCACTTGTGGC 7560



1309 TCACCAATATTCGCGCTCTTTTGGCAATGTGAGGGCCGGAAACCTGGCCCTGCTCTCTTG 1368  
1321 TCACCAATATTCGCGCTCTTTTGGCAATGTGAGGGCCGGAAACCTGGCCCTGCTCTCTTG 1380  
1369 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCAAGAGGAATCAAGGTCTGTTGAATGTC 1428  
1381 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCAAGAGGAATCAAGGTCTGTTGAATGTC 1440  
1429 GTGAAGAGAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACAGTCTGTAGCGACCTT 1488  
1441 GTGAAGAGAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACAGTCTGTAGCGACCTT 1500  
1489 TCGAGCGAGCGAAACCCCCCACTGCGGACAGGTGCTCTGCGGCCAAAGCGACGTGTA 1548  
1501 TCGAGCGAGCGAAACCCCCCACTGCGGACAGGTGCTCTGCGGCCAAAGCGACGTGTA 1560  
1549 TAAAGATACACTGCAAAAGCGGCACAAACCCAGTGCACGTGTGAGTTGGATGTTG 1608  
1561 TAAAGATACACTGCAAAAGCGGCACAAACCCAGTGCACGTGTGAGTTGGATGTTG 1620  
1609 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAGAA 1668  
1621 GAAAGAGTCAAAATGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAGAA 1680  
1669 GTACCCCATTTGATGGGATCTGATCTGGGGCTCTGCGTGCACATGCTTTACATGTGTTAG 1728  
1681 GTACCCCATTTGATGGGATCTGATCTGGGGCTCTGCGTGCACATGCTTTACATGTGTTAG 1740  
1729 TCGAGGTTAAAAAGCTGTAGGCCCCCGAACCAAGGGGACGTGTTTCCCTTTGAAAAA 1788  
1741 TCGAGGTTAAAAAGCTGTAGGCCCCCGAACCAAGGGGACGTGTTTCCCTTTGAAAAA 1800  
1789 CACGATAATACCATGCGGCTATTACGGCTACTCCCAACAGACGCGAGGCTTACTTGC 1848  
1801 CACGATAATACCATGCGGCTATTACGGCTACTCCCAACAGACGCGAGGCTTACTTGC 1860  
1849 TGCATCATCACTAGCTACAGCCCGGACAGGAACAGGTGAGGGAGGTCCAAAGTG 1908  
1861 TGCATCATCACTAGCTACAGCCCGGACAGGAACAGGTGAGGGAGGTCCAAAGTG 1920  
1909 GTCTCCACCGCAACAATCTTTCTGGGACCTGGTCAATGGCGTGTGTGGACTGTC 1968  
1921 GTCTCCACCGCAACAATCTTTCTGGGACCTGGTCAATGGCGTGTGTGGACTGTC 1980  
1969 TATCATGTGCGCGCTCAAGACCTTTGCGGCCCAAGGGCCCAATCAACCAATGTAC 2028  
1981 TATCATGTGCGCGCTCAAGACCTTTGCGGCCCAAGGGCCCAATCAACCAATGTAC 2040  
2029 ACCAATGTGACAGGACCTGTGCGGCTGCAAGCGCCCCCGGGCGGCTTCTTGACA 2088  
2041 ACCAATGTGACAGGACCTGTGCGGCTGCAAGCGCCCCCGGGCGGCTTCTTGACA 2100  
2089 CCATGCACTGTGCGGAGCTGCACTTTACTTGGTCAAGAGGATGCGGATGCTATTCCG 2148  
2101 CCATGCACTGTGCGGAGCTGCACTTTACTTGGTCAAGAGGATGCGGATGCTATTCCG 2160  
2149 GTGCGCGGCGGGGACAGCAGGGGAGCTTCTTCCCGAGGCCGCTCTCTACTTTG 2208  
2161 GTGCGCGGCGGGGACAGCAGGGGAGCTTCTTCCCGAGGCCGCTCTCTACTTTG 2220  
2209 AAGGGCTCTTCGGGGCGGTCCACTGTCTGCCCCCTCGGGGACGCTGTGGGCATTTCCG 2268  
2221 AAGGGCTCTTCGGGGCGGTCCACTGTCTGCCCCCTCGGGGACGCTGTGGGCATTTCCG 2280  
2269 GCTGCGGTGTGACCCGAGGGGTTGGAGCGGTGGACTTTGTACCGCTCGAGTCTATG 2328  
2281 GCTGCGGTGTGACCCGAGGGGTTGGAGCGGTGGACTTTGTACCCGCTCGAGTCTATG 2340  
2329 GAAACCACTATGGGTCCCGGCTTTTCAAGCAAACTCTGCTCCCTCCGGCGGTACCGGAG 2388  
2341 GAAACCACTATGGGTCCCGGCTTTTCAAGCAAACTCTGCTCCCTCCGGCGGTACCGGAG 2400  
2389 ACATTCAGGTGCGCCCATCTACAGCCCTTACTGTGTAGCGGCAAGACATAAGGTGCGG 2448

2401 ACATTCAGGTGCGCCCATCTACGCCCTTACTGTAGCGGCAAGACATAAGGTGCGG 2460  
2449 GCTGCGTATGCAAGCCCAAGGGTATAAGGTGCTTGTCTTGAACCCGCTCGTCCGCGCAC 2508  
2461 GCTGCGTATGCAAGCCCAAGGGTATAAGGTGCTTGTCTTGAACCCGCTCGTCCGCGCAC 2520  
2509 CTAGGTTTCGGGGCGTATATGTCTAAGGCACATGTTATGCAACCTTAAACATCAAGACGG 2568  
2521 CTAGGTTTCGGGGCGTATATGTCTAAGGCACATGTTATGCAACCTTAAACATCAAGACGG 2580  
2569 GTAAAGACCATCACACGGTCCGCCCATCAGCTACTCCACCTATGGCAAGTCTTCTTGCC 2628  
2581 GTAAAGACCATCACACGGTCCGCCCATCAGCTACTCCACCTATGGCAAGTCTTCTTGCC 2640  
2629 GACGGTGGTGTCTCTGGGGGCGCTATGACATATAATATGTGATGAGTGCACCTCAACT 2688  
2641 GACGGTGGTGTCTCTGGGGGCGCTATGACATATAATATGTGATGAGTGCACCTCAACT 2700  
2689 GACTCGAACCACTATCTTGGGCATCGGCACAGTCTTGGAACAAGCGGAGAGCGGTGGAGCG 2748  
2701 GACTCGAACCACTATCTTGGGCATCGGCACAGTCTTGGAACAAGCGGAGAGCGGTGGAGCG 2760  
2749 CGACTCGTCTGCTCGCCACCGCTAGCCTCGGGATCGGTCAACCTGCGCACATCCAAAC 2808  
2761 CGACTCGTCTGCTCGCCACCGCTAGCCTCGGGATCGGTCAACCTGCGCACATCCAAAC 2820  
2809 ATCGAGGAGTGGCTCTGTCCAGCACTGGAGAAATCCCCCTTTATGCGAAAGCAATCCCC 2868  
2821 ATCGAGGAGTGGCTCTGTCCAGCACTGGAGAAATCCCCCTTTATGCGAAAGCAATCCCC 2880  
2869 ATCGAGACCATCAAGGGGGAGGACCTCATTTTCTGCANTTCAAGAAAGAAATGTGAT 2928  
2881 ATCGAGACCATCAAGGGGGAGGACCTCATTTTCTGCCATTTCCAAAGAAAGAAATGTGAT 2940  
2929 GAGCTCGCGCGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTT 2988  
2941 GAGCTCGCGCGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTT 3000  
2989 GATGATTCGCTCATACCACTAGCGGAGAGCTATTTGCTAGCAACGGAAGCTCTTAATG 3048  
3001 GATGATTCGCTCATACCACTAGCGGAGAGCTATTTGCTAGCAACGGAAGCTCTTAATG 3060  
3049 ACGGGCTTTACCGGCGATTTGACTCAGTGATGCTGCAATACATGTGTACCCAGACA 3108  
3061 ACGGGCTTTACCGGCTGACTTCGACTCAGTGATGCTGCAATACATGTGTACCCAGACA 3120  
3109 GTGCACTTCAGCTCGGACCGGCTTACCATTTGAGACGACGACCTGCGCACAAAGACGG 3168  
3121 GTGCACTTCAGCTCGGACCGGCTTACCATTTGAGACGACGACCTGCGCACAAAGACGG 3180  
3169 GTGCACTTCGCGGCGGAGGAGGACTGTGTAGGGGAGGATGGGCAATTTACAGGTTT 3228  
3181 GTGCACTTCGCGGCGGAGGAGGACTGTGTAGGGGAGGATGGGCAATTTACAGGTTT 3240  
3229 GTGACTCCAGGAGAAACGGGCTTCGGGCATGTCGATTCCTCGTTCCTGTGCGAGTGTAT 3288  
3241 GTGACTCCAGGAGAAACGGGCTTCGGGCATGTCGATTCCTCGTTCCTGTGCGAGTGTAT 3300  
3289 GAGCGGGCTGTGCTGTGTACGAGCTCACGCCCGGAGACCTCAGTTAGTTTGGGGCT 3348  
3301 GAGCGGGCTGTGCTGTGTACGAGCTCACGCCCGGAGACCTCAGTTAGTTTGGGGCT 3360  
3349 TACTTAAACACACCGGTTGCCGCTCTGCCAGGACCTCTGAGTTCCTGGGAGACGCTC 3408  
3361 TACTTAAACACACCGGTTGCCGCTCTGCCAGGACCTCTGAGTTCCTGGGAGACGCTC 3420  
3409 TTTTACAGGCTCACCCACATAGACGCCCATTTTCTGTCACAGCTAAGCAGGAGGAGAC 3468  
3421 TTTTACAGGCTCACCCACATAGACGCCCATTTTCTGTCACAGCTAAGCAGGAGGAGAC 3480  
3469 ACTTCCCTTACTGCTGTAGCATACAGGCTACGGTGTGCGGAGGCTCAGGCTCCACCT 3528





QY 5689 CTTCCGATACACCTCCAGGAGGAGGAGCGTTGTCCTGTGAGAACTCTACCGTGCT 5748  
Db |||||  
QY 5701 CCTCGATACACCTTCCAGGAGGAGGAGCGTTGTCCTGTGAGAACTCTACCGTGCT 5760  
Db |||||  
QY 5749 TCTGCTTTGGGGAGCTCGCCAAAGACCTTTCGGCAGCTCCGAATCTGTCGGCCGTCGAC 5808  
Db |||||  
QY 5761 TCTGCTTTGGGGAGCTCGCCACAGACCTTTCGGCAGCTCCGAATCTGTCGGCCGTCGAC 5820  
Db |||||  
QY 5809 AGCGCAGCGGAACCGCTCTCTGACCAAGCCCTCCGACGCGGAGACCGGAGATCCGAC 5868  
Db |||||  
QY 5821 AGCGCAGCGGAACCGCTCTCTGACCAAGCCCTCCGACGCGGAGACCGGAGATCCGAC 5880  
Db |||||  
QY 5869 GTTGAGTGTACTCTCCATGCGCCCTTTGAGGGGAGCGGGGATCCGATCTCAGC 5928  
Db |||||  
QY 5881 GTTGAGTGTACTCTCCATGCGCCCTTTGAGGGGAGCGGGGATCCGATCTCAGC 5940  
Db |||||  
QY 5929 GACGGTCTTGTGCTACCGTAAGCAGGAGGCTAGTGAGGAGCTGCTGCTGCTCGATG 5988  
Db |||||  
QY 5941 GACGGTCTTGTGCTACCGTAAGCAGGAGGCTAGTGAGGAGCTGCTGCTGCTCGATG 6000  
Db |||||  
QY 5989 TCCTACATGACAGGCGCTCTCTGACCAAGCCCTCCGACGCGGAGACCGGAGATCCGAC 6048  
Db |||||  
QY 6001 TCCTACATGACAGGCGCTCTCTGACCAAGCCCTCCGACGCGGAGACCGGAGATCCGAC 6060  
Db |||||  
QY 6049 ATCAATGCACTGAGCACTCTTGTGCTCGCTCACCACAACTTGTGCTATGCTACCAATCT 6108  
Db |||||  
QY 6061 ATCAATGCACTGAGCACTCTTGTGCTCGCTCACCACAACTTGTGCTATGCTACCAATCT 6120  
Db |||||  
QY 6109 CGCAGCGCAAGCTCGCGGAGAGAGGTCACCTTTGACAGACTGCAAGTCTCTGAGACGAC 6168  
Db |||||  
QY 6121 CGCAGCGCAAGCTCGCGGAGAGAGGTCACCTTTGACAGACTGCAAGTCTCTGAGACGAC 6180  
Db |||||  
QY 6169 CACTACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTAAGGCTAAACTT 6228  
Db |||||  
QY 6181 CACTACCGGAGCGTCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTAAGGCTAAACTT 6240  
Db |||||  
QY 6229 CTATCGTGAGGAGCGCTTGAAGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6288  
Db |||||  
QY 6241 CTATCGTGAGGAGCGCTTGAAGTGAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6300  
Db |||||  
QY 6289 TATGGGCAAGAGCTCGCGGAGCGCTTACAGCAAGCGGCTTAAACACATCGCTCCGCTG 6348  
Db |||||  
QY 6301 TATGGGCAAGAGCTCGCGGAGCGCTTACAGCAAGCGGCTTAAACACATCGCTCCGCTG 6360  
Db |||||  
QY 6349 TGAAGGACTTGTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGCGCAAAAT 6408  
Db |||||  
QY 6361 TGAAGGACTTGTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGCGCAAAAT 6420  
Db |||||  
QY 6409 GAGTTTCTGCTGCTCAACAGAGAGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6468  
Db |||||  
QY 6421 GAGTTTCTGCTGCTCAACAGAGAGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6480  
Db |||||  
QY 6469 CCAGATTGTTGGGCTTGTGTCGAGAAATGGCGCTTTACGATGTGCTCCACCCCTC 6528  
Db |||||  
QY 6481 CCAGATTGTTGGGCTTGTGTCGAGAAATGGCGCTTTACGATGTGCTCCACCCCTC 6540  
Db |||||  
QY 6529 CCTCAGCGCTGATGGCTTCTCATCAGGATTCATATCTCTCCTGAGACAGCGGCTCGAG 6588  
Db |||||  
QY 6541 CCTCAGCGCTGATGGCTTCTCATCAGGATTCATATCTCTCCTGAGACAGCGGCTCGAG 6600  
Db |||||  
QY 6589 TTCTGTTGATGCTTGAAGCGGAGAAATGGCTTATGGCTTCGCATATGACACCCG 6648  
Db |||||  
QY 6601 TTCTGTTGATGCTTGAAGCGGAGAAATGGCTTATGGCTTCGCATATGACACCCG 6660  
Db |||||  
QY 6649 TGTGTTGACTCAACGGTCACTGAGATGACATCGGTGTTGAGGAGTCAATCTACCAATGT 6708  
Db |||||  
QY 6661 TGTGTTGACTCAACGGTCACTGAGATGACATCGGTGTTGAGGAGTCAATCTACCAATGT 6720  
Db |||||  
QY 6709 TGTGACTTGGCCCCGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6768  
Db |||||  
QY 6721 TGTGACTTGGCCCCGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 6780  
Db |||||  
QY 6769 GGGGGCCCCCTGACTAAATCTAAGGCGAGAACTGCGGCTATCGCGGTGCGCGGCGGAGC 6828

Db |||||  
QY 6781 GGGGGCCCCCTGACTAAATCTAAGGCGAGAACTGCGGCTATCGCGGTGCGCGGAGC 6840  
Db |||||  
QY 6829 GGTGTACTGACGACAGCTGCGGTAAATACCTTCACATGTTACTTGAAGCCGCTCGCGGC 6888  
Db |||||  
QY 6841 GGTGTACTGACGACAGCTGCGGTAAATACCTTCACATGTTACTTGAAGCCGCTCGCGGC 6900  
Db |||||  
QY 6889 TGTGAGCTGCGAAGCTCCAGACTGCGAGCTGCTGATGCGGAGACCACTTGTGCTT 6948  
Db |||||  
QY 6901 TGTGAGCTGCGAAGCTCCAGACTGCGAGTGTCTGATGCGGAGACCACTTGTGCTT 6960  
Db |||||  
QY 6949 ATCTGTGAAAGCGCGGAGACCAAGAGGAGGAGGCTACGGGCTTACGGAGGCT 7008  
Db |||||  
QY 6961 ATCTGTGAAAGCGCGGAGACCAAGAGGAGGAGGCTACGGGCTTACGGAGGCT 7020  
Db |||||  
QY 7009 ATGACTAGATCTCTGCGCCCTGGGACCGCGCAAAACCAAGATACGATTTGAGTTG 7068  
Db |||||  
QY 7021 ATGACTAGATCTCTGCGCCCTGGGACCGCGCAAAACCAAGATACGATTTGAGTTG 7080  
Db |||||  
QY 7069 ATAAATCATGCTCTCTCAATGTGTGCTGCGCAGGATGATCTGGCAAAAGGTTGATC 7128  
Db |||||  
QY 7081 ATAAATCATGCTCTCTCAATGTGTGCTGCGCAGGATGATCTGGCAAAAGGTTGATC 7140  
Db |||||  
QY 7129 TATCTCACCGCTGACCCACACCCCTTGGCGGCTGCTGGAGACAGCTAGACAC 7188  
Db |||||  
QY 7141 TATCTCACCGCTGACCCACACCCCTTGGCGGCTGCTGGAGACAGCTAGACAC 7200  
Db |||||  
QY 7189 ACTCCAGTCAATCTCTGCTAGGCAACATCATGTATGCGCCCACTTGTGGGCAAG 7248  
Db |||||  
QY 7201 ACTCCAGTCAATCTCTGCTAGGCAACATCATGTATGCGCCCACTTGTGGGCAAG 7260  
Db |||||  
QY 7249 ATGATCTGATGACTCATTTCTTCTCATCTTCTAGCTCAGGAACTTGAAGAGCC 7308  
Db |||||  
QY 7261 ATGATCTGATGACTCATTTCTTCTCATCTTCTAGCTCAGGAACTTGAAGAGCC 7320  
Db |||||  
QY 7309 CTAGATTGTGAGATCTACGGGCTGTTACTCCTATGAGCCACTTACCTCAGATC 7368  
Db |||||  
QY 7321 CTAGATTGTGAGATCTACGGGCTGTTACTCCTATGAGCCACTTACCTCAGATC 7380  
Db |||||  
QY 7369 ATTCAACGACTCCATGCGCTTAGCGCAATTTCACTCCATGATTTCTCCAGTGTAGATC 7428  
Db |||||  
QY 7381 ATTCAACGACTCCATGCGCTTAGCGCAATTTCACTCCATGATTTACTCTCAGGTGAGATC 7440  
Db |||||  
QY 7429 AATAGGTTGCTTCACTGCTCAGGAACTTGGGTACCGCCCTTGGAGTCTGGAGACAT 7488  
Db |||||  
QY 7441 AATAGGTTGCTTCACTGCTCAGGAACTTGGGTACCGCCCTTGGAGTCTGGAGACAT 7500  
Db |||||  
QY 7489 CGGGCCAGAAAGTGTCCGCTTAGGCTTACTGTCCAGGGGGGAGGCTGCCACTTGTGGC 7548  
Db |||||  
QY 7501 CGGGCCAGAAAGTGTCCGCTTAGGCTTACTGTCCAGGGGGGAGGCTGCCACTTGTGGC 7560  
Db |||||  
QY 7549 AAGTACTCTTCAACTGGGAGTAAGGACCAAGCTCAAACTCACTCAATFCCCGGCTGCG 7608  
Db |||||  
QY 7561 AAGTACTCTTCAACTGGGAGTAAGGACCAAGCTCAAACTCACTCAATFCCCGGCTGCG 7620  
Db |||||  
QY 7609 TCCAGTCTGATTTATCAGCTGCTTGTGCTGTTTACGCGGGGAGACATATAC 7668  
Db |||||  
QY 7621 TCCAGTCTGATTTATCAGCTGCTTGTGCTGTTTACGCGGGGAGACATATAC 7680  
Db |||||  
QY 7669 AGCTGCTCTGTCGCGGACCGCTGTTTATGTTGCTTACTCTCTCTTCTGTAGGG 7728  
Db |||||  
QY 7681 AGCTGCTCTGTCGCGGACCGCTGTTTATGTTGCTTACTCTCTCTTCTGTAGGG 7740  
Db |||||  
QY 7729 GTAGGATCTATCTACTCCCCAACCGATGAACGGGACCTAAACACTCCAGGCAATAGG 7788  
Db |||||  
QY 7741 GTAGGATCTATCTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCAATAGG 7800  
Db |||||  
QY 7789 CAACTGCTTTTTTCCCTTT 7848  
Db |||||  
QY 7801 CAACTGCTTTTTTCCCTTT 7860  
Db |||||  
QY 7849 TTTTCTCTTTTTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTAGC 7908



QY 1609 GAAAGAGTCAATGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGATGCCAGAG 1668  
DB 1621 GAAAGAGTCAATGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGATGCCAGAG 1680  
QY 1669 GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTGTTAG 1728  
DB 1681 GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTGTTAG 1740  
QY 1729 TCGAGGTTAAAAACGCTTAGGCCCCCGAACCAACGAGGACGTGTTTCTTTGAAAAA 1788  
DB 1741 TCGAGGTTAAAAACGCTTAGGCCCCCGAACCAACGAGGACGTGTTTCTTTGAAAAA 1800  
QY 1789 CAGGATATACCATGGCGCTATTAGCGCTACTCTCCACAGACGCGAGCGCTACTTGGC 1848  
DB 1801 CAGGATATACCATGGCGCTATTAGCGCTACTCTCCACAGAGCGAGCGCTACTTGGC 1860  
QY 1849 TGCATCATCACTAGCTACAGGCGGGGACAGGAACCAAGTCAAGGGGAGGTCCAAGTG 1908  
DB 1861 TGCATCATCACTAGCTACAGGCGGGGACAGGAACCAAGTCAAGGGGAGGTCCAAGTG 1920  
QY 1909 GTCTCCACGCAACAATCTTTCTGGCGACTGCTGCTCAATGGCGGTGTTGGACTGTC 1968  
DB 1921 GTCTCCACGCAACAATCTTTCTGGCGACTGCTGCTCAATGGCGGTGTTGGACTGTC 1980  
QY 1969 TATCATGTCGGGCTCAAGACCTTTCGGCGCCCAAGGCGCGGCTTCTTTGACA 2028  
DB 1981 TATCATGTCGGGCTCAAGACCTTTCGGCGCCCAAGGCGCGGCTTCTTTGACA 2040  
QY 2029 ACCAATGTGACACAGGACCTCTGCGTGGCGAGCGCCCGGGCGGCTTCTTTGACA 2088  
DB 2041 ACCAATGTGACACAGGACCTCTGCGTGGCGAGCGCCCGGGCGGCTTCTTTGACA 2100  
QY 2089 CCAATGCACTGCGGCGAGCTCGGACCTTTACTTGGTCAAGAGGATCGGATGTCATCCG 2148  
DB 2101 CCAATGCACTGCGGCGAGCTCGGACCTTTACTTGGTCAAGAGGATCGGATGTCATCCG 2160  
QY 2149 GTGCGCGGGGGGACAGAGGGGAGCTACTCTCCCGAGGCGGCTCTCTACTTGG 2208  
DB 2161 GTGCGCGGGGGGACAGAGGGGAGCTACTCTCCCGAGGCGGCTCTCTACTTGG 2220  
QY 2209 AAGGGCTCTTGGGCGGTCCACTGCTCTGCCCTCGGGGACGCTCTGGGCACTTTTCG 2268  
DB 2221 AAGGGCTCTTGGGCGGTCCACTGCTCTGCCCTCGGGGACGCTCTGGGCACTTTTCG 2280  
QY 2269 GTGCGGTGTCACCGAGGGGTTCGAAAGGCGGTGGAATTTGTAACCGTFCGAGTCTATG 2328  
DB 2281 GTGCGGTGTCACCGAGGGGTTCGAAAGGCGGTGGAATTTGTAACCGTFCGAGTCTATG 2340  
QY 2329 GAAACCACTATGGTCCCGGTCTTACGGCAACTCGTCCCTCGGCGCGTACCGCAG 2388  
DB 2341 GAAACCACTATGGTCCCGGTCTTACGGCAACTCGTCCCTCGGCGCGTACCGCAG 2400  
QY 2389 ACATTCAGGTGCCCATCTACACGCGCCCTACTGTGTAGCGGCAAGCACTAAGGTGCCG 2448  
DB 2401 ACATTCAGGTGCCCATCTACACGCGCCCTACTGTGTAGCGGCAAGCACTAAGGTGCCG 2460  
QY 2449 GTGCGGTATGACGCCCAAGGATATAGGTGCTTGTCTCTGAAACCGTCTGCGCGCGAAC 2508  
DB 2461 GTGCGGTATGACGCCCAAGGATATAGGTGCTTGTCTCTGAAACCGTCTGCGCGCGAAC 2520  
QY 2509 CTAGGTTTCGGGCGGTATATGCTTAGGGCACATGTTATCGACCTTAACATCAACCGG 2568  
DB 2521 CTAGGTTTCGGGCGGTATATGCTTAGGGCACATGTTATCGACCTTAACATCAACCGG 2580  
QY 2569 GTAAGGACCATCACCGGTGCGCCCATCACGTACTTCCACCTATGGCAAGTTTCTTGGC 2628  
DB 2581 GTAAGGACCATCACCGGTGCGCCCATCACGTACTTCCACCTATGGCAAGTTTCTTGGC 2640  
QY 2629 GACGGTGTGCTCTGGGCGGCTTATGACATCATATATGATGATGATGATGATGATCAACT 2688  
DB 2641 GACGGTGTGCTCTGGGCGGCTTATGACATCATATATGATGATGATGATGATGATCAACT 2700  
QY 2689 GACTCGACCACTATCTTGGGCATCGGCACAGTCTCTGGAACCAAGGGAGAGCGGTGGAG 2748

DB 2701 GACTCGACCACTATCTTGGGCATCGGCACAGTCTTGGACCAAGCGGAGAGCGGTGGAGG 2760  
QY 2749 CGACTCGTCTGCTCGCCACCGCTACGCTCCGGATCGGTACCGTGCACATCCAAAC 2808  
DB 2761 CGACTCGTCTGCTCGCCACCGCTACGCTCCGGATCGGTACCGTGCACATCCAAAC 2820  
QY 2809 ATCGAGGAGGTGGCTCTGTCTAGACTGAGAGAAATCCCTTTTATGCGAAAGCATCCCC 2868  
DB 2821 ATCGAGGAGGTGGCTCTGTCTAGACTGAGAGAAATCCCTTTTATGCGAAAGCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGAGGACCTCATTTCTGCAATTCGAAGAAATGTGAT 2928  
DB 2881 ATCGAGACCATCAAGGGGGAGGACCTCATTTCTGCAATTCGAAGAAATGTGAT 2940  
QY 2929 GAGCTCGCCCGGAAGCTGTCCGGCCCTCGGACTCAATGTGTAGCATATTTACCGGGGCTT 2988  
DB 2941 GAGCTCGCCCGGAAGCTGTCCGGCCCTCGGACTCAATGTGTAGCATATTTACCGGGGCTT 3000  
QY 2989 GATGATCTCGTATACCAACTAGCGGAGAGCTCATTTCTGTAGCAACGAGCGCTCTAATG 3048  
DB 3001 GATGATCTCGTATACCAACTAGCGGAGAGCTCATTTCTGTAGCAACGAGCGCTCTAATG 3060  
QY 3049 ACGGGCTTTACCGGCGATTTCGACTCAGTGATCGACTGCAATACATGTGTCAACAGACA 3108  
DB 3061 ACGGGCTTTACCGGCGATTTCGACTCAGTGATCGACTGCAATACATGTGTCAACAGACA 3120  
QY 3109 GTCGACTTCAGCTCGGACCGGCTTACCATGTAGAGACGACCGCTGCGCAAGACGG 3168  
DB 3121 GTCGACTTCAGCTCGGACCGGCTTACCATGTAGAGACGACCGCTGCGCAAGACGG 3180  
QY 3169 GTGTCACGCTTCGAGCGGGGAGGAGGACTGTGTAGGGGAGGATGGGATTTACAGTTT 3228  
DB 3181 GTGTCACGCTTCGAGCGGGGAGGAGGACTGTGTAGGGGAGGATGGGATTTACAGTTT 3240  
QY 3229 GTGACTTCAGGAGAAACGGGCTTCGGCATGTTCTCGTTCCTGTCGAGTGTCTAT 3288  
DB 3241 GTGACTTCAGGAGAAACGGGCTTCGGCATGTTCTCGTTCCTGTCGAGTGTCTAT 3300  
QY 3289 GACGGGGCTGTCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTGCGGGCT 3348  
DB 3301 GACGGGGCTGTCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTGCGGGCT 3360  
QY 3349 TACTTAAACACACGAGGTTCGCTGCGAGGACCATCTGGAGTTCTGGGAGAGCGTC 3408  
DB 3361 TACTTAAACACACGAGGTTCGCTGCGAGGACCATCTGGAGTTCTGGGAGAGCGTC 3420  
QY 3409 TTTACAGGCTCACCCACATAGACGCCCATTTCTTCTCCAGACTAAGCAGGCGAGAGAC 3468  
DB 3421 TTTACAGGCTCACCCACATAGACGCCCATTTCTTCTCCAGACTAAGCAGGCGAGAGAC 3480  
QY 3469 AACTTCCCCTACTGTGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCT 3528  
DB 3481 AACTTCCCCTACTGTGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCT 3540  
QY 3529 CCATGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAGGCTACGCTGACCGGCCA 3588  
DB 3541 CCATGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAGGCTACGCTGACCGGCCA 3600  
QY 3589 AGCGCCCTCTGTATAGGCTGGGAGCGGTTTAAACAGAGGTTACTACCAACACCCCAT 3648  
DB 3601 AGCGCCCTCTGTATAGGCTGGGAGCGGTTTAAACAGAGGTTACTACCAACACCCCAT 3660  
QY 3649 ACCAATACATCATGSCATGCTGCTGAGCTGAGGCTGCTACAGAGCACTCGGGTG 3708  
DB 3661 ACCAATACATCATGSCATGCTGCTGAGCTGAGGCTGCTACAGAGCACTCGGGTG 3720  
QY 3709 CTGTTAGGCGGAGTCTTACAGCTCTGGCGGGTATTGCTGTACAAACAGGAGCGGTGTC 3768  
DB 3721 CTGTTAGGCGGAGTCTTACAGCTCTGGCGGGTATTGCTGTACAAACAGGAGCGGTGTC 3780  
QY 3769 ATTGTGGGAGGATCATCTTGTCCGAAAGCGGCCCATCATTTCCGACAGGAAGTCTT 3828

Db 3781 ATTGTGGCAGGATCATCTTGTCCGAAAGCCGGCCATCATTTCCCGACAGGGAAATCTCTT 3840  
QY 3829 TACCCGGAGTTTCATGAGATGGAAGAGTGGCTCACACCTCCCTTAATCGAAAGGGA 3888  
Db 3841 TACCCGGAGTTTCATGAGATGGAAGAGTGGCTCACACCTCCCTTAATCGAAAGGGA 3900  
QY 3889 ATGCAGCTCGCCGAAACAAATTCAGACAGAGGCAATCGGGTTCCTGCAACAGCCACCAG 3948  
Db 3901 ATGCAGCTCGCCGAAACAAATTCAGACAGAGGCAATCGGGTTCCTGCAACAGCCACCAG 3960  
QY 3949 CAAGCGAGGCTGTCTCCCTGGTGGAAATCAAGTGGCGGACCCCTCGAAGCTTCTGG 4008  
Db 3961 CAAGCGAGGCTGTCTCCCTGGTGGAAATCAAGTGGCGGACCCATCGAAGCTTCTGG 4020  
QY 4009 GCGAGCATATGTGAAATTCATCAGCGGGATACAAATTTAGCAGGCTTGTCCACTGTG 4068  
Db 4021 GCGAGCATATGTGAAATTCATCAGCGGGATACAAATTTAGCAGGCTTGTCCACTGTG 4080  
QY 4069 CCTGCCAACCCCGCATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCCCGCTC 4128  
Db 4081 CCTGCCAACCCCGCATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCCCGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTGTTTAAATCTCTGGGGGATGGGTGGCCGCCAACTTCT 4188  
Db 4141 ACCACCAACATACCTCTCTGTTTAAATCTCTGGGGGATGGGTGGCCGCCAACTTCT 4200  
QY 4189 CCTCCAGCGCTTCTCTTCTGCTTCTAGCGCGGCATGCTGGAGCGGCTGTTGGCAGC 4248  
Db 4201 CCTCCAGCGCTTCTCTTCTGCTTCTAGCGCGGCATGCTGGAGCGGCTGTTGGCAGC 4260  
QY 4249 ATAGGCTTTGGGAAGGCTGTGTGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGC 4308  
Db 4261 ATAGGCTTTGGGAAGGCTGTGTGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGC 4320  
QY 4309 GCGTCTGGGCTTTAAGGTATAGCGCGGAGATGCCCTCCACCGAGAACCTTGTAAAC 4368  
Db 4321 GCGTCTGGGCTTTAAGGTATAGCGCGGAGATGCCCTCCACCGAGAACCTTGTAAAC 4380  
QY 4369 CTACTCCCTGTATCTCTCCCTGGCGCCTAGTCTGTCGGGTCGTGTGGCAGCAT 4428  
Db 4381 CTACTCCCTGTATCTCTCCCTGGCGCCTAGTCTGTCGGGTCGTGTGGCAGCAT 4440  
QY 4429 CTGGCTCGGCACGTGGGCGCAGGGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCG 4488  
Db 4441 CTGGCTCGGCACGTGGGCGCAGGGGAGGGGCTGTGCAGTGGATGAACCGGCTGATAGCG 4500  
QY 4489 TTGGCTTCGGGGGTAAACAGTCTCCCTCCAGCATATGTCCCTGAGAGGACGCTGCA 4548  
Db 4501 TTGGCTTCGGGGGTAAACAGTCTCCCTCCAGCATATGTCCCTGAGAGGACGCTGCA 4560  
QY 4549 GCAGTGTCACTCAGATCTCTTAGTCTTACCATCACTCAGCTGCTGCAAGAGGCTTCCAC 4608  
Db 4561 GCAGTGTCACTCAGATCTCTTAGTCTTACCATCACTCAGCTGCTGAAAGGCTTCCAC 4620  
QY 4609 CAGTGGATCAACGAGGACTGTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGG 4668  
Db 4621 CAGTGGATCAACGAGGACTGTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGG 4680  
QY 4669 GATTGGATATGACAGGCTTGACTGATTTCAAGACCTGGCTGCAGTCCAGCTCTGCGG 4728  
Db 4681 GATTGGATATGACAGGCTTGACTGATTTCAAGACCTGGCTGCAGTCCAGTCCAGCTCTGCGG 4740  
QY 4729 CGATTGCCGGAGTCCCTCTTCTCTCATGTCAAAGTGGGTACAAAGGAGTCTGGCGGGGC 4788  
Db 4741 CGATTGCCGGAGTCCCTCTTCTCTCATGTCAAAGTGGGTACAAAGGAGTCTGGCGGGGC 4800  
QY 4789 GACGGCATATGCAAAACCACTTCCATGTGGAGCACAGATCACCGGACATGTGAAAAAC 4848  
Db 4801 GACGGCATATGCAAAACCACTTCCATGTGGGGCACAGATCACCGGACATGTGAAAAAC 4860  
QY 4849 GGTTCATGAGGATCGTGGGCTTAGGACCTGTAGTAACAGTGGCATGGAACATTTCCC 4908  
Db 4861 GGTTCATGAGGATCGTGGGCTTAGGACCTGTAGTAACAGTGGCATGGAACATTTCCC 4920

QY 4909 ATTAACGGTATACACACGGGCCCCGTGACAGCCCTCCCGGCCCAAAATTTATTTCTAGGGCG 4968  
Db 4921 ATTAACGGTATACACACGGGCCCCGTGACAGCCCTCCCGGCCCAAAATTTATTTCTAGGGCG 4980  
QY 4969 CTGTGGCGGGTGGCTGTGAGGAGTACGTGAGAGTTTACGCGGTGGGGATTTCCACTAC 5028  
Db 4981 CTGTGGCGGGTGGCTGTGAGGAGTACGTGAGAGTTTACGCGGTGGGGATTTCCACTAC 5040  
QY 5029 GTGACGGGCATGACCACTGACAAAGTAAAGTGCCTGTGAGGTTCCGSCCCCGAAATTC 5088  
Db 5041 GTGACGGGCATGACCACTGACAAAGTAAAGTGCCTGTGAGGTTCCGSCCCCGAAATTC 5100  
QY 5089 TTCAAGAGTGGATGGGGTGGGTGACAGGTACGTTCAGGTGCAAGCTGCAAAACCCCTCTTA 5148  
Db 5101 TTCAAGAGTGGATGGGGTGGGTGACAGGTACGTTCAGGTGCAAGCTGCAAAACCCCTCTTA 5160  
QY 5149 CGGAGGAGGTCAATCTCGTGGGTCAATCAATACCTGTTGGGTACAGCTCCCA 5208  
Db 5161 CGGAGGAGGTCAATCTCGTGGGTCAATCAATACCTGTTGGGTACAGCTCCCA 5220  
QY 5209 TCGAGGCCGAAACCGGACGTAGCTACTTCCATGCTCACGACCCCTCCCAATTC 5268  
Db 5221 TCGAGGCCGAAACCGGACGTAGCTACTTCCATGCTCACGACCCCTCCCAATTC 5280  
QY 5269 ACGGCGGAGCGGCTAAGGCTAGGTAGGCGCAGGGATCTCCCTCTTGGCCAGCTCA 5328  
Db 5281 ACGGCGGAGCGGCTAAGGCTAGGTAGGCGCAGGGATCTCCCTCTTGGCCAGCTCA 5340  
QY 5329 TCAGCTAGCAGCTGTCTGCGCTTCTTGAAGGCAACATGACATACCTGTCATCAGCTCC 5388  
Db 5341 TCAGCTAGCAGCTGTCTGCGCTTCTTGAAGGCAACATGACATACCTGTCATCAGCTCC 5400  
QY 5389 CCGAGCGTGACCTCATCGAGGCCAACCTCTGTGGCGCAGGAGATGGCGGGAAATTC 5448  
Db 5401 CCGAGCGTGACCTCATCGAGGCCAACCTCTGTGGCGCAGGAGATGGCGGGAAATTC 5460  
QY 5449 ACCCGGTGGAGTCAGAAATAAGGTAGTAAATTTGGACTCTTTCGAGCCGCTCCAGCG 5508  
Db 5461 ACCCGGTGGAGTCAGAAATAAGGTAGTAAATTTGGACTCTTTCGAGCCGCTCCAGCG 5520  
QY 5509 GAGGAGATGAGAGGAGTATCGGTTCCGGCGAGATCTCTGCGGAGGTCAGAGAAATTC 5568  
Db 5521 GAGGAGATGAGAGGAGTATCGGTTCCGGCGGAGATCTCTGCGGAGGTCAGAGAAATTC 5580  
QY 5569 CCTCAGGATGCGCATATGGGACGCGCGGATTAACACCTTCCACTTTAGAGTCCTGG 5628  
Db 5581 CCTCAGGATGCGCATATGGGACGCGCGGATTAACACCTTCCACTTTAGAGTCCTGG 5640  
QY 5629 AAGGACCGGACTACGTCCTCCAGTGTACACGGGTGTCATTCGCGCTGCCAAGGCC 5688  
Db 5641 AAGGACCGGACTACGTCCTCCAGTGTACACGGGTGTCATTCGCGCTGCCAAGGCC 5700  
QY 5689 CCTCCGATACCACTCCACGGAGGAGGAGTGTGTCTGTGAGAAATCTACCGTGTCT 5748  
Db 5701 CCTCCGATACCACTCCACGGAGGAGGAGTGTGTCTGTGAGAAATCTACCGTGTCT 5760  
QY 5749 TCTGCTTGGCGAGCTGCGCACAAAGACCTTCGGGAGCTCCGAATCGTGGCGCTGAC 5808  
Db 5761 TCTGCTTGGCGAGCTGCGCACAGAGACCTTCGGGAGCTCCGAATCGTGGCGCTGAC 5820  
QY 5809 AGCGGACGGCAACGGCTCTCTGACAGCCCTCCGACGACGGGAGCGGGATCCGAC 5868  
Db 5821 AGCGGACGGCAACGGCTCTCTGACAGCCCTCCGACGACGGGAGCGGGATCCGAC 5880  
QY 5869 GTTCAGTCTACTCTCTCCATGSCCCCTTTAGGGGAGCGCGGGGATTCGATCTCAGC 5928  
Db 5881 GTTCAGTCTACTCTCTCCATGSCCCCTTTAGGGGAGCGCGGGGATTCGATCTCAGC 5940  
QY 5929 GACGGCTTGTGCTTACCGTAAGCGAGGAGCTAGTGAGGAGCTGCTGCTGCTCGATG 5988  
Db 5941 GACGGCTTGTGCTTACCGTAAGCGAGGAGCTAGTGAGGAGCTGCTGCTGCTCGATG 6000

QY 5989 TCCTACATGACAGGCGCCCTGATCAGCCATGCGTGGGAGGAACCAAGCTGCC 6048  
DB 6001 TCTTACATGACAGGCGCCCTGATCAGCCATGCGTGGGAGGAACCAAGCTGCC 6060  
QY 6049 ATCAATGCACTGAGCAACTCTTTGCTCCGCTCACCAACTTGGTCTTATCTACAACTCT 6108  
DB 6061 ATCAATGCACTGAGCAACTCTTTGCTCCGCTCACCAACTTGGTCTTATCTACAACTCT 6120  
QY 6109 CGCAGCGCAAGCCTCGGCGAGAAAGGTCACTTTGACAGACTGCAAGTCTCTGAGAGAC 6168  
DB 6121 CGCAGCGCAAGCCTCGGCGAGAAAGGTCACTTTGACAGACTGCAAGTCTCTGAGAGAC 6180  
QY 6169 CACTACCGGAGCGTCTCAAGAGATGAAGCGAAGCGCTCCACAGTTAAGGTTAACTT 6228  
DB 6181 CACTACCGGAGCGTCTCAAGAGATGAAGCGAAGCGCTCCACAGTTAAGGTTAACTT 6240  
QY 6229 CTATCCGTGGAGAGCCTGTAAAGCTGACGCCGCCACATTCGGCCAGATCTAAATTTGGC 6288  
DB 6241 CTATCCGTGGAGAGCCTGTAAAGCTGACGCCGCCACATTCGGCCAGATCTAAATTTGGC 6300  
QY 6289 TATGGGCAAAAGACTCGTCGGAACCTTATCCAGCAAGCGGTTAAACCAATCCGCTCCGTG 6348  
DB 6301 TATGGGCAAAAGACTCGTCGGAACCTTATCCAGCAAGCGGTTAAACCAATCCGCTCCGTG 6360  
QY 6349 TGGAGGACTTCTGCGTCCAAACGAGAGAGGGGGCGCAAGCCAGCTCGCCTTATCGTATTTC 6408  
DB 6361 TGGAGGACTTCTGCGTCCAAACGAGAGACTGAGACCAAAATTGACACCAACCATCATGGCAAAAT 6420  
QY 6409 GAGGTTTCTGCGTCCAAACGAGAGAGGGGGCGCAAGCCAGCTCGCCTTATCGTATTTC 6468  
DB 6421 GAGGTTTCTGCGTCCAAACGAGAGAGGGGGCGCAAGCCAGCTCGCCTTATCGTATTTC 6480  
QY 6469 CCAGATTTGGGGTTCGTGTGCGAGAAATAGGCCCTTTACATGATGTGCTCCACCTTC 6528  
DB 6481 CCAGATTTGGGGTTCGTGTGCGAGAAATAGGCCCTTTACATGATGTGCTCCACCTTC 6540  
QY 6529 CCTCAGGCGCTGATGGCTCTTCATACGATTCCTAATCTCTCGAGAGCGGGTCGAG 6588  
DB 6541 CCTCAGGCGCTGATGGCTCTTCATACGATTCCTAATCTCTCGAGAGCGGGTCGAG 6600  
QY 6589 TTCTCGTGAATGCCCTGGAAGCGAAGAAATGCCCTATGGGCTTCGCATATGACACCCGC 6648  
DB 6601 TTCTCGTGAATGCCCTGGAAGCGAAGAAATGCCCTATGGGCTTCGCATATGACACCCGC 6660  
QY 6649 TGTGTTGACTCAACGGTCTACTGAGATGATCATCCGTTGTGAGGAGTCAATCTACCAATGT 6708  
DB 6661 TGTGTTGACTCAACGGTCTACTGAGATGATCATCCGTTGTGAGGAGTCAATCTACCAATGT 6720  
QY 6709 TGTGACTTGGCCCCGAAAGCGACAGAGGCGCATTAAGGTTCGCTACAGAGCGGCTTTACATC 6768  
DB 6721 TGTGACTTGGCCCCGAAAGCGACAGAGGCGCATTAAGGTTCGCTACAGAGCGGCTTTACATC 6780  
QY 6769 GGGGGCCCCCTGACTAAATTTAAAGGCGCAAACTGGGGCTATCGCCGGTCCCGCGCGAGC 6828  
DB 6781 GGGGGCCCCCTGACTAAATTTAAAGGCGCAAACTGGGGCTATCGCCGGTCCCGCGCGAGC 6840  
QY 6829 GGTGTACTGAGCAACAGCTGGGTATATACCTCACAATGTTACTGTAAGGCGGCTCGGCC 6888  
DB 6841 GGTGTACTGAGCAACAGCTGGGTATATACCTCACAATGTTACTGTAAGGCGGCTCGGCC 6900  
QY 6889 TGTGAGCTGCGAGACTCCAGGACTCAGAGTCTCGTATGGGAGCGACCTTGTGCTT 6948  
DB 6901 TGTGAGCTGCGAGACTCCAGGACTCAGAGTCTCGTATGGGAGCGACCTTGTGCTT 6960  
QY 6949 ATCTGTGAAGCGCGGAGCCCAAGAGGACGAGCGGCGCTTACGGGCTTACGGAGGCT 7008  
DB 6961 ATCTGTGAAGCGCGGAGCCCAAGAGGACGAGCGGCGCTTACGGGCTTACGGAGGCT 7020  
QY 7009 ATGACTAGATPACTCTGCCCCCTGCGGAGCCCGCCCAACAGAAATGAGTGTGAGTTG 7068  
DB 7021 ATGACTAGATPACTCTGCCCCCTGCGGAGCCCGCCCAACAGAAATGAGTGTGAGTTG 7080  
QY 7069 ATAACATCATGCTCTCCCAATGTGTAGTCCGCGCAGCATCTGGCAAAAGGGTGTAC 7128

DB 7081 ATPAACATCATGCTCTCTCAATGTGTAGTCCGCGCAGATGCACTGGCAAAAGGGTGTAC 7140  
QY 7129 TATCTACCCCGTGACCCCAACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC 7188  
DB 7141 TATCTACCCCGTGACCCCAACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACAC 7200  
QY 7189 ACTCAGTCAATCTCTGCTAGGCAACATCATCATGTATGCGCCCACTTGTGGGCAAG 7248  
DB 7201 ACTCAGTCAATCTCTGCTAGGCAACATCATCATGTATGCGCCCACTTGTGGGCAAG 7260  
QY 7249 ATGATCTCTGATGACTCATTTCTTCCATCTTCTAGCTCAGGAACTCTGAAAAAGCC 7308  
DB 7261 ATGATCTCTGATGACTCATTTCTTCCATCTTCTAGCTCAGGAACTCTGAAAAAGCC 7320  
QY 7309 CTAGATTTGTCAGATCTAGCGGCGCTTTACTCCATGAGCCACTTGACCTCAGATC 7368  
DB 7321 CTAGATTTGTCAGATCTAGCGGCGCTTTACTCCATGAGCCACTTGACCTCAGATC 7380  
QY 7369 ATTCAACGACTCCATGGCCTTAGCGATTTTCACTCCATAGTTACTCTCAGGTGAGATC 7428  
DB 7381 ATTCAACGACTCCATGGCCTTAGCGATTTTCACTCCATAGTTACTCTCAGGTGAGATC 7440  
QY 7429 AATAGGTTGGCTTCATGCTCAGGAAACTTTGGGGTACCGCCCTTGGAGTCTGGAGACAT 7488  
DB 7441 AATAGGTTGGCTTCATGCTCAGGAAACTTTGGGGTACCGCCCTTGGAGTCTGGAGACAT 7500  
QY 7489 CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGC 7548  
DB 7501 CGGGCCAGAAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGC 7560  
QY 7549 AAGTACCTCTTCAACTGGGCACTAGGACCAAGCTCCTCACTCCATCCCGGCTGG 7608  
DB 7561 AAGTACCTCTTCAACTGGGCACTAGGACCAAGCTCCTCACTCCATCCCGGCTGG 7620  
QY 7609 TCCAGTTGGATTTATCCAGCTGGTTCGTGCTGTTACAGGGGGGAGACATATATCAC 7668  
DB 7621 TCCAGTTGGATTTATCCAGCTGGTTCGTGCTGTTACAGGGGGGAGACATATATCAC 7680  
QY 7669 AGCTGTCTCTGCGCGACCCCGCTGTTTCACTGTTGCTACTCTCTCTTCTGTAGGG 7728  
DB 7681 AGCTGTCTCTGCGCGACCCCGCTGTTTCACTGTTGCTACTCTCTCTTCTGTAGGG 7740  
QY 7729 GTAGGACTATCTACTCTCCCAACCGATGAAAGGGGACCTTAACACTCCAGGCCATAGG 7788  
DB 7741 GTAGGACTATCTACTCTCCCAACCGATGAAAGGGGACCTTAACACTCCAGGCCATAGG 7800  
QY 7789 CCATCTCTGTTTTTTCCTT 7848  
DB 7801 CCATCTCTGTTTTTTCCTT 7860  
QY 7849 TTTTCTCTTTTTTTTCTCTTTTTTCTTTTTTCTTTTTTCTTTTCTTTTCTTTTCTTTT 7908  
DB 7861 TTTTCTCTTTTTTTTCTCTTTTTTCTTTTTTCTTTTTTCTTTTCTTTTCTTTTCTTTT 7920  
QY 7909 CCTAGTCAGCGTACTGTGAAGGTCCTGAGCCGCTTGACTGACAGAGTCTCTATAC 7968  
DB 7921 CCTAGTCAGCGTACTGTGAAGGTCCTGAGCCGCTTGACTGACAGAGTCTCTATAC 7980  
QY 7969 TGGCCTCTCTGAGATCAAGT 7989  
DB 7981 TGGCCTCTCTGAGATCAAGT 8001

## RESULT 10

AR406045  
LOCUS AR406045 8001 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 16 from patent US 6630343.  
ACCESSION AR406045  
VERSION AR406045.1 GI:40155172  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

Unclassified.	
REFERENCE 1 (bases 1 to 8001)	
AUTHORS Bartenschlager,R.	
TITLE Hepatitis C virus culture system	
JOURNAL Patent: US 6630343-A 16 07-OCT-2003;	
FEATURES Location/Qualifiers	
source 1..8001	
/organism="unknown"	
/mol_type="genomic DNA"	
ORIGIN	
Query Match 99.5%; Score 7946.2; DB 6; Length 8001;	
Best Local Similarity 99.7%; Pred. No. 0;	
Matches 7976; Conservative 0; Mismatches 13; Indels 12; Gaps 1;	
QY 1 GCCAGCCCCGATGGGGGCGACATCCACCATAGATCACTCCCCCTGTGAGGAACACTCTG 60	
DB 1 GCCAGCCCCGATGGGGGCGACATCCACCATAGATCACTCCCCCTGTGAGGAACACTCTG 60	
QY 61 TCTTCACGCAGAAAGCGTAGCCATAGTGGTCTGGGGAACCGGTGAGTACACCGGAATTGCCAG 120	
DB 61 TCTTCACGCAGAAAGCGTAGCCATAGTGGTCTGGGGAACCGGTGAGTACACCGGAATTGCCAG 120	
QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGGGGAACCGGTGAGTACACCGGAATTGCCAG 180	
DB 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGGGGAACCGGTGAGTACACCGGAATTGCCAG 180	
QY 181 GACGACCGGTCTCTTCTTGGATCAACCCGCTCAATGCTCGAGATTGGCGCGCCGCC 240	
DB 181 GACGACCGGTCTCTTCTTGGATCAACCCGCTCAATGCTCGAGATTGGCGCGTCCGCC 240	
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGTCTGGGGAACCGGTGAGTACACCGGAATTGCCAG 300	
DB 241 GCGAGACTGTAGCCGAGTAGTGTGGTCTGGGGAACCGGTGAGTACACCGGAATTGCCAG 300	
QY 301 GTGCTTGCAGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCTTAAC 360	
DB 301 GTGCTTGCAGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCTTAAC 360	
QY 361 CTCGAAGAAACCAAA-----GGGCGGCCATGATTGAACAAGATGATGC 408	
DB 361 CTCGAAGAAACCAAAACGTAAACACCGGGCGCCCATGATTGAACAAGATGATGC 420	
QY 409 ACGCAGGTTCTCCGGCGCTTGGTGGAGAGGTATTGGTACTGCTACCTCCCTGATAG 468	
DB 421 ACGCAGGTTCTCCGGCGCTTGGTGGAGAGGTATTGGTACTGCTACCTCCCTGATAG 480	
QY 469 CAATCGGCTGCTGTATGCCCGCTGTTCCGGCTGTGAGCGCAGGGGCGCCGGTCTTT 528	
DB 481 CAATCGGCTGCTGTATGCCCGCTGTTCCGGCTGTGAGCGCAGGGGCGCCGGTCTTT 540	
QY 529 TTGTCAAGACCGACTGTCGGTCCCTGAATGAATGACATGACATGAGCAGCGCGGTAT 588	
DB 541 TTGTCAAGACCGACTGTCGGTCCCTGAATGAATGACATGACATGAGCAGCGCGGTAT 600	
QY 589 CGTGGCTGGCCACACCGGGCGTCTCTTGGCAGCTGTGCTGACGTTGTCACTCAAGCG 648	
DB 601 CGTGGCTGGCCACACCGGGCGTCTCTTGGCAGCTGTGCTGACGTTGTCACTCAAGCG 660	
QY 649 GAAGGACTGGCTGCTATTGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACCTTG 708	
DB 661 GAAGGACTGGCTGCTATTGGCGAAGTGGCGGGCAGGATCTCTGTCTCATCTCACCTTG 720	
QY 709 CTCCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCGGGCGCTGCATACGCTTATC 768	
DB 721 CTCCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCGGGCGCTGCATACGCTTATC 780	
QY 769 CGGTACTCTGCCATTCGACCAACAAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGA 828	
DB 781 CGGTACTCTGCCATTCGACCAACAAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGA 840	
QY 829 TGGNAGCGGTCTTGTGATCAGATGATCTGGACGAAGAGCATCAGGGCTCCGCCAG 888	
DB 841 TGGNAGCGGTCTTGTGATCAGATGATCTGGACGAAGAGCATCAGGGCTCCGCCAG 900	
QY 889 CCGAACTGTTCCCGAGGCTCAAGCGCGCATGCCCGACGGGAGGATCTCGTGTGACCC 948	
DB 901 CCGAACTGTTCCCGAGGCTCAAGCGCGCATGCCCGACGGGAGGATCTCGTGTGACCC 960	
QY 949 ATGGCGATGCTGCTTGGCGAATATCATGTTGGAAATATGGCGGCTTTCTGGATTATCG 1008	
DB 961 ATGGCGATGCTGCTTGGCGAATATCATGTTGGAAATATGGCGGCTTTCTGGATTATCG 1020	
QY 1009 ACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGACATAGCGTTGCTACCGCTGATA 1068	
DB 1021 ACTGTGGCGGCTGGGTGTGGCGACCGCTATCAGACATAGCGTTGCTACCGCTGATA 1080	
QY 1069 TTGTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGCTTTTACGSTATCGCG 1128	
DB 1081 TTGTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCTCTGCTTTTACGSTATCGCG 1140	
QY 1129 CTCCCGATTCCGAGCGGATCGCTTCTATCGCTTCTTGACAGTTCTTCTGAGTTTAAA 1188	
DB 1141 CTCCCGATTCCGAGCGGATCGCTTCTATCGCTTCTTGACAGTTCTTCTGAGTTTAAA 1200	
QY 1189 CAGACCAACACGCTTCCCTCTAGCGGATCAATTCGGCCCTCTCCCTCCCCCCCCCT 1248	
DB 1201 CAGACCAACACGCTTCCCTCTAGCGGATCAATTCGGCCCTCTCCCTCCCCCCCCCT 1260	
QY 1249 AACCTTACTGCCGAAAGCGCTTGGAAATAGGCGGTGTGGTGTGTATATGTTATTT 1308	
DB 1261 AACCTTACTGCCGAAAGCGCTTGGAAATAGGCGGTGTGGTGTGTATATGTTATTT 1320	
QY 1309 TCCACCATATTGCGCTCTTTTGGCAATGAGGGCGCGGAAACCTTGGGCCCTCTCTCTG 1368	
DB 1321 TCCACCATATTGCGCTCTTTTGGCAATGAGGGCGCGGAAACCTTGGGCCCTCTCTCTG 1380	
QY 1369 ACGAGCATTTCTAGGGGTCTTTTCCCTCTCCGCAAAAGGAATGCAAGTCTGTGAATGTC 1428	
DB 1381 ACGAGCATTTCTAGGGGTCTTTTCCCTCTCCGCAAAAGGAATGCAAGTCTGTGAATGTC 1440	
QY 1429 GTGAAGAGCAGTTCTCTGGAAGCTTTTGAAGACAAACCAACCTGTGTAGCACCTTT 1488	
DB 1441 GTGAAGAGCAGTTCTCTGGAAGCTTTTGAAGACAAACCAACCTGTGTAGCACCTTT 1500	
QY 1489 TGCAGGACAGGGAACCCCGACCTGGCGACAGGTCTCTCGGCGCAAAAGCAGCTGTA 1548	
DB 1501 TGCAGGACAGGGAACCCCGACCTGGCGACAGGTCTCTCGGCGCAAAAGCAGCTGTA 1560	
QY 1549 TAAGATACCTGCAAAAGGCGGCAACCCCGACAGTGTGAGTTGATAGTTGTG 1608	
DB 1561 TAAGATACCTGCAAAAGGCGGCAACCCCGACAGTGTGAGTTGATAGTTGTG 1620	
QY 1609 GAAAGATCAAAATGGCTCTCTCAAGCGTATTCAAAGGGGCTGAAGGATGCCAGAG 1668	
DB 1621 GAAAGATCAAAATGGCTCTCTCAAGCGTATTCAAAGGGGCTGAAGGATGCCAGAG 1680	
QY 1669 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCAATCTTTACATGTTTGTAG 1728	
DB 1681 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCAATCTTTACATGTTTGTAG 1740	
QY 1729 TCGAGGTTAAAAACGCTAGGGCCCCCGAACCAAGGGGCGTGGTCTTCTTTTGAATAA 1788	
DB 1741 TCGAGGTTAAAAACGCTAGGGCCCCCGAACCAAGGGGCGTGGTCTTCTTTTGAATAA 1800	
QY 1789 CAGATTAATACCATGGCGCTATTACGGCTATCTCCCAACAGACGCGAGGCTACTTGGC 1848	
DB 1801 CAGATTAATACCATGGCGCTATTACGGCTATCTCCCAACAGACGCGAGGCTACTTGGC 1860	
QY 1849 TGCATCATCACTAGCTCAGCGCGGACAGGAACAGGTCTGAGGGGGAGGTCACAGTG 1908	
DB 1861 TGCATCATCACTAGCTCAGCGCGGACAGGAACAGGTCTGAGGGGGAGGTCACAGTG 1920	
QY 1909 GTCTCCACCGCAACAAATCTTTCTGGCGACCTTGGCTCAATGGCGTGTGTGACTGTC 1968	
DB 1921 GTCTCCACCGCAACAAATCTTTCTGGCGACCTTGGCTCAATGGCGTGTGTGACTGTC 1980	



QY 1969 TATCATGTGCGGCTCAAAAGACCTTTCGGCCCAAAAGGGGCCCAATCAACCCAAATGTAC 2028  
Db 1981 TATCATGTGCGGCTCAAAAGACCTTTCGGCCCAAAAGGGGCCCAATCAACCCAAATGTAC 2040  
QY 2029 ACCAATGTGGACAGGACTCTGTGGCTGGCAAGCGCCCCCGGGCGGTTTCCTTGACA 2088  
Db 2041 ACCAATGTGGACAGGACTCTGTGGCTGGCAAGCGCCCCCGGGCGGTTTCCTTGACA 2100  
QY 2089 CCATGCACCTGTGGCAGCTCGGACCTTTACTTTGGTCAAGAGCATGCCGATCTCATTCG 2148  
Db 2101 CCATGCACCTGTGGCAGCTCGGACCTTTACTTTGGTCAAGAGCATGCCGATCTCATTCG 2160  
QY 2149 GTGGCCCGGGGGGCGACAGAGGGGAGCCTACTCTCCCGAGCGCGTCTCTACTTG 2208  
Db 2161 GTGGCCCGGGGGGCGACAGAGGGGAGCCTACTCTCCCGAGCGCGTCTCTACTTG 2220  
QY 2209 AAGGCTCTTTCGGGGGCTCCACTGTCTGCCCCCTCGGGGCAAGCTGTGGGCATCTTTCGG 2268  
Db 2221 AAGGCTCTTTCGGGGGCTCCACTGTCTGCCCCCTCGGGGCAAGCTGTGGGCATCTTTCGG 2280  
QY 2269 GGTGCGGTGTGCACCGAGGGGTTTCGAAGGGCGGTGGAATTTGTATACCGTTCGAGTCTATG 2328  
Db 2281 GGTGCGGTGTGCACCGAGGGGTTTCGAAGGGCGGTGGAATTTGTATACCGTTCGAGTCTATG 2340  
QY 2329 GAAACCACTATGCGGTGCCCGTCTTACGGAACAAGTCTGCTCCCTCGGCGGCTACCGGAG 2388  
Db 2341 GAAACCACTATGCGGTGCCCGTCTTACGGAACAAGTCTGCTCCCTCGGCGGCTACCGGAG 2400  
QY 2389 ACATTTCCAGGTGGCCCATCTACAGCCCTACTGTGTAGCGGCAAGAGCACTAAGGTGCG 2448  
Db 2401 ACATTTCCAGGTGGCCCATCTACAGCCCTACTGTGTAGCGGCAAGAGCACTAAGGTGCG 2460  
QY 2449 GGTGCGGTATGCAGCCCAAGGGTATAAGGTGCTTGTCTGAAACCGGTCGGTCCGCGCCACC 2508  
Db 2461 GGTGCGGTATGCAGCCCAAGGGTATAAGGTGCTTGTCTGAAACCGGTCGGTCCGCGCCACC 2520  
QY 2509 CTAGGTTTCGGGGGCTATATGTCTAAGGCACATGTGTATCGACCTAAACATCAGAACCGGG 2568  
Db 2521 CTAGGTTTCGGGGGCTATATGTCTAAGGCACATGTGTATCGACCTAAACATCAGAACCGGG 2580  
QY 2569 GTAAGGACCATCACACGGGTCGCCCATCAGTACTCCACTATGCGCAAGTTCCTTGC 2628  
Db 2581 GTAGGACCATCACACGGGTCGCCCATCAGTACTCCACTATGCGCAAGTTCCTTGC 2640  
QY 2629 GACGTTGTTGCTCTGGGGGCGCTATGACATATAATATGTATGATGAGTGCCACTCAACT 2688  
Db 2641 GACGTTGTTGCTCTGGGGGCGCTATGACATATAATATGTATGATGAGTGCCACTCAACT 2700  
QY 2689 GACTCGACCACTATCTTGGGATCGGCACAGTCTCGGACCAAGCGGAGCGCTGGAGCG 2748  
Db 2701 GACTCGACCACTATCTTGGGATCGGCACAGTCTCGGACCAAGCGGAGCGCTGGAGCG 2760  
QY 2749 CCACTCGTGTCTGCGCACCGCTACGCTCCGGGATCGGTCAAGTGCACATCCAAAC 2808  
Db 2761 CCACTCGTGTCTGCGCACCGCTACGCTCCGGGATCGGTCAAGTGCACATCCAAAC 2820  
QY 2809 ATCGAGGAGGTGGCTGTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCC 2868  
Db 2821 ATCGAGGAGGTGGCTGTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGGAGGCACTCAATTTCTGCAATTCMAAGAAATGTGAT 2928  
Db 2881 ATCGGACCATCAAGGGGGGAGGCACTCAATTTCTGCAATTCMAAGAAATGTGAT 2940  
QY 2929 GAGTCTCGCGGAAGCTGTCCGCTCGGACTCAATGCTGTAGCATATTACGGGCGCTT 2988  
Db 2941 GAGTCTCGCGGAAGCTGTCCGCTCGGACTCAATGCTGTAGCATATTACGGGCGCTT 3000  
QY 2989 GATGTATCGGTATACCAACTAGCGGAGAGCTCATTTGTCTGTAGCAACGAGCGCTCTAATG 3048  
Db 3001 GATGTATCGGTATACCAACTAGCGGAGAGCTCATTTGTCTGTAGCAACGAGCGCTCTAATG 3060

QY 3049 ACGGCTTTTACGGGGATTTTCGACTCAGTGATCGACTGCAATACATGTGTCAACCCAGACA 3108  
Db 3061 ACGGCTTTTACGGGGATTTTCGACTCAGTGATCGACTGCAATACATGTGTCAACCCAGACA 3120  
QY 3109 GTGCACTTTCAGCTTGCAGCCGACCTTACCACTTTGAGAGCGACACCGTGCACAAAGCGG 3168  
Db 3121 GTGCACTTTCAGCTTGCAGCCGACCTTACCACTTTGAGAGCGACACCGTGCACAAAGCGG 3180  
QY 3169 GTGTCACTCTCGAGCGGCGAGGAGGAGTCTGTGTAGGGGAGGATGGGCATTTACAGTTT 3228  
Db 3181 GTGTCACTCTCGAGCGGCGAGGAGGAGTCTGTGTAGGGGAGGATGGGCATTTACAGTTT 3240  
QY 3229 GTGACTCCAGAGAGAACCGGCTTCGGGCAATGTTCGATTCCTCGGTTCTGTGCGAGTGTAT 3288  
Db 3241 GTGACTCCAGAGAGAACCGGCTTCGGGCAATGTTCGATTCCTCGGTTCTGTGCGAGTGTAT 3300  
QY 3289 GACGCGGCTGTCTGTGTGTAGAGCTACGCGCCCGGAGACCTCAGTTAGTGTTCGGGCT 3348  
Db 3301 GACGCGGCTGTCTGTGTGTAGAGCTACGCGCCCGGAGACCTCAGTTAGTGTTCGGGCT 3360  
QY 3349 TACTTAAACACACCGAGGTTCGCTTCGCGAGGACCATCTGTGAGTTCGTGGAGAGCGTC 3408  
Db 3361 TACTTAAACACACCGAGGTTCGCTTCGCGAGGACCATCTGTGAGTTCGTGGAGAGCGTC 3420  
QY 3409 TTTACAGGCTTACCCACATAGACGCCATTTCTTGTCCAGACTAAGAGCGAGAGAC 3468  
Db 3421 TTTACAGGCTTACCCACATAGACGCCATTTCTTGTCCAGACTAAGAGCGAGAGAC 3480  
QY 3469 AACTTCCCTTACTCTGTAGCATACAGGCTACGCTGTGGCGGAGGCTCAGGCTCCACT 3528  
Db 3481 AACTTCCCTTACTCTGTGTAGCATACAGGCTACGCTGTGGCGGAGGCTCAGGCTCCACT 3540  
QY 3529 CCATCGTGGGACCAATGTGGAAGTGTCTCATACGCTTAAAGCTTACCGTGCACGGGCA 3588  
Db 3541 CCATCGTGGGACCAATGTGGAAGTGTCTCATACGCTTAAAGCTTACCGTGCACGGGCA 3600  
QY 3589 AGCGGCTGTGTATAGCTGGAGCGGTTCAAAAGAGGTTACTACACACACCCCAT 3648  
Db 3601 AGCGGCTGTGTATAGCTGGAGCGGTTCAAAAGAGGTTACTACACACACCCCAT 3660  
QY 3649 ACCAAATACATCATGGCATGCAATGTGGCTGACCTGAGGTCGTCAAGAGCTTGGGTG 3708  
Db 3661 ACCAAATACATCATGGCATGCAATGTGGCTGACCTGAGGTCGTCAAGAGCTTGGGTG 3720  
QY 3709 CTGTGTAGCGGAGTCTGTAGAGCTGTGGCGGCTATGCTGCAACAGGAGCGGTGTC 3768  
Db 3721 CTGTGTAGCGGAGTCTGTAGAGCTGTGGCGGCTATGCTGCAACAGGAGCGGTGTC 3780  
QY 3769 ATTGTGGCGAGGATCATCTGTTCGGAAGCGGCTTCCATTCGAGAGGAAAGTCTT 3828  
Db 3781 ATTGTGGCGAGGATCATCTGTTCGGAAGCGGCTTCCATTCGAGAGGAAAGTCTT 3840  
QY 3829 TACCGGAGTTCGATGAGATGGAAGTGGCTCACACTCCCTTACATCGAAACAGGGA 3888  
Db 3841 TACCGGAGTTCGATGAGATGGAAGTGGCTCACACTCCCTTACATCGAAACAGGGA 3900  
QY 3889 ATGCACTCGCGCAACAAATTCAAAAGAGGCAATTCGGGTTGCTGCAACAGGAGCGGTC 3948  
Db 3901 ATGCACTCGCGCAACAAATTCAAAAGAGGCAATTCGGGTTGCTGCAACAGGAGCGGTC 3960  
QY 3949 GAAAGCGGAGTGTCTGCTCCGTTGGAAATCCAAAGTGGGAGCCCTCGAAGCGTCTG 4008  
Db 3961 GAAAGCGGAGTGTCTGCTCCGTTGGAAATCCAAAGTGGGAGCCCTCGAAGCGTCTG 4020  
QY 4009 GCGAAGCATATGTGGAATTTTCAATCAGCGGATACAAATTTTAGCAGGCTTGTCCACTCTG 4068  
Db 4021 GCGAAGCATATGTGGAATTTTCAATCAGCGGATACAAATTTTAGCAGGCTTGTCCACTCTG 4080  
QY 4069 CTGCAACCCCGGAGTACATCTGATGGCATTCACAGCTCTATACAGCGCGCTC 4128  
Db 4081 CTGCAACCCCGGAGTACATCTGATGGCATTCACAGCTCTATACAGCGCGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTGTTTAAATCTCTGGGGGATGGTGGCCCGCACTTGTCT 4188

Db 4141 ACCACCCAAACATACCTCTCTTTAAACATCTTGGGGGATGGGTGGCGCCCAACTTGCT 4200  
Qy CTCTCCAGCGTCTTCTGCTTTTGTAGGCGCGCATCGCTGGAGCGGCTGTGGCAGC 4248  
Db 4201 CCTCCAGCGTCTTCTGCTTTTGTAGGCGCGCATCGCTGGAGCGGCTGTGGCAGC 4260  
Qy 4249 ATAGGCTTTGGGAAGTCTTGTGGATATTTTGGCAGGTATGGAGCAGGGGTGGCAGC 4308  
Db 4261 ATAGGCTTTGGGAAGTCTTGTGGATATTTTGGCAGGTATGGAGCAGGGGTGGCAGC 4320  
Qy 4309 GCGTCTGTGGCTTTAAGGTCAAGAGCGCGAGATGCCCTCCACCGAGAACCTGGTTAAC 4368  
Db 4321 GCGTCTGTGGCTTTAAGGTCAAGAGCGCGAGATGCCCTCCACCGAGAACCTGGTTAAC 4380  
Qy 4369 CTACTCCCTGTATCTCTCCCTGGCGCCCTAGTCGTCGGGGTCTGTGGCAGCGATA 4428  
Db 4381 CTACTCCCTGTATCTCTCCCTGGCGCCCTAGTCGTCGGGGTCTGTGGCAGCGATA 4440  
Qy 4429 CTGCGTCGGCACGTGGGCCCAAGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGC 4488  
Db 4441 CTGCGTCGGCACGTGGGCCCAAGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGC 4500  
Qy 4489 TTGCGTTCGGGGTAAACAGTCTTCCCGCAGCGACATATGTGCGCTGAGAGCGACGTGCA 4548  
Db 4501 TTGCGTTCGGGGTAAACAGTCTTCCCGCAGCGACATATGTGCGCTGAGAGCGACGTGCA 4560  
Qy 4549 GCAGTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGCTTCAC 4608  
Db 4561 GCAGTGTCACTCAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGCTTCAC 4620  
Qy 4609 CAGTGGATCAACAGGAGTGTCCACGCCATGCTCCGGCTCGTGCTGAAGAGTGTGG 4668  
Db 4621 CAGTGGATCAACAGGAGTGTCCACGCCATGCTCCGGCTCGTGCTGAAGAGTGTGG 4680  
Qy 4669 GATTGGATATGCAAGGTGTGACTGATTTTCAAGACCTGGGCTCAGTCAAGCTCTGCG 4728  
Db 4681 GATTGGATATGCAAGGTGTGACTGATTTTCAAGACCTGGGCTCAGTCAAGCTCTGCG 4740  
Qy 4729 CGATTGCGGGAGTCCCTCTCTCTATGTCACAGTGGGTACAAAGGAGTCTGGCGGGC 4788  
Db 4741 CGATTGCGGGAGTCCCTCTCTCTATGTCACAGTGGGTACAAAGGAGTCTGGCGGGC 4800  
Qy 4789 GACGGCATATGCAAAACACCTGCCATGTGGAGCAGAGTACCGGACATGTGAAAAAC 4848  
Db 4801 GACGGCATATGCAAAACACCTGCCATGTGGAGCAGAGTACCGGACATGTGAAAAAC 4860  
Qy 4849 GGTTCATAGAGATCGTGGGCTTAGACCTGTAGTAAACAGTGGCATGGAACTTCCCC 4908  
Db 4861 GGTTCATAGAGATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACTTCCCC 4920  
Qy 4909 ATTAAGCGGTACACACGGGCCCTGACAGCCCTCCCGCGCCAAATTTCTAGGGCG 4968  
Db 4921 ATTAAGCGGTACACACGGGCCCTGACAGCCCTCCCGCGCCAAATTTCTAGGGCG 4980  
Qy 4969 CTGTGGGGGTGCTGTAGAGTACGTGGAGTTACGGGGTGGGGATTTCCACTAC 5028  
Db 4981 CTGTGGGGGTGCTGTAGAGTACGTGGAGTTACGGGGTGGGGATTTCCACTAC 5040  
Qy 5029 GTGACGGCATGACCACTGACAAAGTAAAGTCCCGCTGTCAGTTCCGGCCCCCGAATTC 5088  
Db 5041 GTGACGGCATGACCACTGACAAAGTAAAGTCCCGCTGTCAGTTCCGGCCCCCGAATTC 5100  
Qy 5089 TTCAAGAGTGGATGGGGTGGGTGACAGGTACGCTCCAGGCTGCAAAACCTCTCTA 5148  
Db 5101 TTCAAGAGTGGATGGGGTGGGTGACAGGTACGCTCCAGGCTGCAAAACCTCTCTA 5160  
Qy 5149 CGGAGAGGTGCAATCTCTGGTCCGGCTCAATCAATACCTGTGGGTACAGCTCCCA 5208  
Db 5161 CGGAGAGGTGCAATCTCTGGTCCGGCTCAATCAATACCTGTGGGTACAGCTCCCA 5220  
Qy 5209 TGGAGCCGAACCGGAGTAGCAGTGTCTACTTCCATGCTCAGCGACCTCCCACTT 5268

Db 5221 TGGGGGCCGAACCGGACGTAGCAGTGTCTACTTCCATGCTCACCGACCCCTCCACACTT 5280  
Qy 5269 ACGGCGGAGACGGCTAAGCGTAGGCTGGCAGGGGATCTCCCTCTCTTGGCAGCTCA 5328  
Db 5281 ACGGCGGAGACGGCTAAGCGTAGGCTGGCAGGGGATCTCCCTCTCTTGGCAGCTCA 5340  
Qy 5329 TCAGTACGACAGTGTCTGCGCTTCTTTGAAGGCAACATCACTACCCGTCATCACTCC 5388  
Db 5341 TCAGTACGACAGTGTCTGCGCTTCTTTGAAGGCAACATCACTACCCGTCATCACTCC 5400  
Qy 5389 CCGAGCGTACCTCATCGAGGCCAACCTCTCTGGCGGAGAGATGGCGGGAACATC 5448  
Db 5401 CCGAGCGTACCTCATCGAGGCCAACCTCTCTGGCGGAGAGATGGCGGGAACATC 5460  
Qy 5449 ACCCGCTGGAGTCAGAAATTAAGTAGTATTTTGGACTCTTTTCGAGCCCTCCAAAGC 5508  
Db 5461 ACCCGCTGGAGTCAGAAATTAAGTAGTATTTTGGACTCTTTTCGAGCCCTCCAAAGC 5520  
Qy 5509 GAGGAGATGAGAGGGAGTATCCGTTCCGGCGAGATCTCTGGGAGGTCCAGGAAATTC 5568  
Db 5521 GAGGAGATGAGAGGGAGTATCCGTTCCGGCGAGATCTCTGGGAGGTCCAGGAAATTC 5580  
Qy 5569 CCTCGAGCATGCGCATATAGGCGACGCCCGGANTPACAACTCCACTGTGTAGAGTCTGG 5628  
Db 5581 CCTCGAGCATGCGCATATAGGCGACGCCCGGANTPACAACTCCACTGTGTAGAGTCTGG 5640  
Qy 5629 AAGGACCGGACTACGTCCTCCAGTGTACACGGGTGTCATTTGCCGCTGCCAAGGCC 5688  
Db 5641 AAGGACCGGACTACGTCCTCCAGTGTACACGGGTGTCATTTGCCGCTGCCAAGGCC 5700  
Qy 5689 CCTCCGATACACCTCCAGGAGAGGACGGTTGTCTGTTCAGATCTACCGTGTCT 5748  
Db 5701 CCTCCGATACACCTCCAGGAGGAGGAGCGGTTGTCTGTTCAGATCTACCGTGTCT 5760  
Qy 5749 TCTGCTTGGCGAGTCTGCCCAAAAGACCTTCCGAGTCTCGGCGGTGAC 5808  
Db 5761 TCTGCTTGGCGAGTCTGCCCAAAAGACCTTCCGAGTCTCGGCGGTGAC 5820  
Qy 5809 AGCGGACGCAACGGCTCTCTGACAGCCCTCCGAGCGGAGCGGAGTCCGAC 5868  
Db 5821 AGCGGACGCAACGGCTCTCTGACAGCCCTCCGAGCGGAGCGGAGTCCGAC 5880  
Qy 5869 GTTGAGTCTGACTCTCTCCATGCCCCCTTGGGGGAGCGGGGATCCCGATCTCAGC 5928  
Db 5881 GTTGAGTCTGACTCTCTCCATGCCCCCTTGGGGGAGCGGGGATCCCGATCTCAGC 5940  
Qy 5929 GACGGTCTTGGTCTACCGTAAGCGAGAGGCTAGTGAAGACGTCGTCTGTCTCGATG 5988  
Db 5941 GACGGTCTTGGTCTACCGTAAGCGAGAGGCTAGTGAAGACGTCGTCTGTCTCGATG 6000  
Qy 5989 TCCTACACATGGACAGCGCCCTGATCACGCCATGCGCTGGGAGGAAACCAAGCTGCC 6048  
Db 6001 TCCTACACATGGACAGCGCCCTGATCACGCCATGCGCTGGGAGGAAACCAAGCTGCC 6060  
Qy 6049 ATCAATCACTGAGCAACTCTTTGCTCCGTCACCACTTTGGTCTATGTCTACCACTCT 6108  
Db 6061 GTCAATCACTGAGCAACTCTTTGCTCCGTCACCACTTTGGTCTATGTCTACCACTCT 6120  
Qy 6109 CGAGCGCAAGCCTGCGGAGAGAGAGTCACTTTGACAGACTGAGAGTCTGGAGCAGC 6168  
Db 6121 CGAGCGCAAGCCTGCGGAGAGAGAGTCACTTTTACAGACTGAGAGTCTGGAGCAGC 6180  
Qy 6169 CACTACCGGAGCTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTTAAGGCTAACTT 6228  
Db 6181 CACTACCGGAGCTGCTCAAGGAGATGAAGCGGAGCGTCCACAGTTAAGGCTAACTT 6240  
Qy 6229 CTATCCGTGGAGGAGCCTGTAAAGTGAACCCCCCATTTGGCCAGATCTAAATTTGGC 6288  
Db 6241 CTATCCGTGGAGGAGCCTGTAAAGTGAACCCCCCATTTGGCCAGATCTAAATTTGGC 6300  
Qy 6289 TATGGGCAAGAGCCTGCGGAACCTTCCAGAAAGCGGCTTAAACCATCCGCTCCGTG 6348  
Db 6301 TATGGGCAAGAGCCTGCGGAACCTTCCAGAAAGCGGCTTAAACCATCCGCTCCGTG 6360

```
QY 6349 TGGAAAGACTTGCTCGAAGACACTGAGACACAAATGACACCAATCATATGCAACAAAAT 6408
Db 6361 TGGAAAGACTTGCTCGAAGACACTGAGACACAAATGACACCAATCATATGCAACAAAAT 6420
QY 6409 GAGGTTTTCTGCGTCCAAACAGAGAGAGGGGGCCGCAAGCCAGCTCGCCTATTCGTATTC 6468
Db 6421 GAGGTTTTCTGCGTCCAAACAGAGAGAGGGGGCCGCAAGCCAGCTCGCCTATTCGTATTC 6480
QY 6469 CCAGATTTGGGGGTTGCGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTC 6528
Db 6481 CCAGATTTGGGGGTTGCGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTC 6540
QY 6529 CCTCAGGCGGTGATGGGCTTCTTACCGAATTCGAATCTCTCTGGAACAGCGGGTCGAG 6588
Db 6541 CCTCAGGCGGTGATGGGCTTCTTACCGAATTCGAATCTCTCTGGAACAGCGGGTCGAG 6600
QY 6589 TTCTCTGGTGAATGCTGGAAGCGGAAGAAATGCCCTATGCGGCTTCGCATATGACACCCG 6648
Db 6601 TTCTCTGGTGAATGCTGGAAGCGGAAGAAATGCCCTATGCGGCTTCGCATATGACACCCG 6660
QY 6649 TGTCTTGTACTCAACGCTCACTGAGATGACATCCGTTGTTGAGGAGTCAATCTACCAATGT 6708
Db 6661 TGTCTTGTACTCAACGCTCACTGAGATGACATCCGTTGTTGAGGAGTCAATCTACCAATGT 6720
QY 6709 TGTGACTTGGCCCGGAGCGAGACGAGGCAATAGGTCTCTCAGAGCGGCTTTACATC 6768
Db 6721 TGTGACTTGGCCCGGAGCGAGACGAGGCAATAGGTCTCTCAGAGCGGCTTTACATC 6780
QY 6769 GGGGGCCCCCTGACTAATTTCTAAAGGCGAGACTCGGCTATGCGGGTCCGCGGAGC 6828
Db 6781 GGGGGCCCCCTGACTAATTTCTAAAGGCGAGACTCGGCTATGCGGGTCCGCGGAGC 6840
QY 6829 GGTGTACTGACGACCACTGCGGTAAATACCCTCACATGTTACTTGAAGGCGGCTCGCGCC 6888
Db 6841 GGTGTACTGACGACCACTGCGGTAAATACCCTCACATGTTACTTGAAGGCGGCTCGCGCC 6900
QY 6889 TGTGAGCTGCGAAGCTCAGGACTGCAAGTCTCGTATGCGGAGACGACCTTGTGTT 6948
Db 6901 TGTGAGCTGCGAAGCTCAGGACTGCAAGTCTCGTATGCGGAGACGACCTTGTGTT 6960
QY 6949 ATCTGTGAAGCGGGGACCCAGAGGACGAGGCGCTACGGGCTTCAAGGAGCT 7008
Db 6961 ATCTGTGAAGCGGGGACCCAGAGGACGAGGCGCTACGGGCTTCAAGGAGCT 7020
QY 7009 ATGACTAGATCTCTGCCCGGAGCCGCGGACCCGCAACCAAGATACGACTTGGAGTTG 7068
Db 7021 ATGACTAGATCTCTGCCCGGAGCCGCGGACCCGCAACCAAGATACGACTTGGAGTTG 7080
QY 7069 ATAAATCATGCTCTCCAAATGTGTCAGTCGCGACGATGCAATCTGGAAAGGGTGTAC 7128
Db 7081 ATAAATCATGCTCTCCAAATGTGTCAGTCGCGACGATGCAATCTGGAAAGGGTGTAC 7140
QY 7129 TATCTCACCGTGACCCACCCCTTGTGCGGGCTGCGTGGGAGACAGTACACAC 7188
Db 7141 TATCTCACCGTGACCCACCCCTTGTGCGGGCTGCGTGGGAGACAGTACACAC 7200
QY 7189 ACTCCAGTCAATTCCTGGCTAGGCAACATCATGATGTCGCCCACTTGTGGGCAAGG 7248
Db 7201 ACTCCAGTCAATTCCTGGCTAGGCAACATCATGATGTCGCCCACTTGTGGGCAAGG 7260
QY 7249 ATGATCCTGATGACTCATTTCTCTCATCTTCTAGCTCAGGAAACAACTTGAAAGGCC 7308
Db 7261 ATGATCCTGATGACTCATTTCTCTCATCTTCTAGCTCAGGAAACAACTTGAAAGGCC 7320
QY 7309 CTAGATTTGACATCTACGGGCGTGTACTCCATTGAGCACTTGCCTACCTCAGATC 7368
Db 7321 CTAGATTTGACATCTACGGGCGTGTACTCCATTGAGCACTTGCCTACCTCAGATC 7380
QY 7369 ATTCAACGACTCCATGGCCCTTAGCGCATTTTCTACTCCATGTTACTCTCCAGGTGAGATC 7428
Db 7381 ATTCAACGACTCCATGGCCCTTAGCGCATTTTCTACTCCATGTTACTCTCCAGGTGAGATC 7440
```

```
QY 7429 AATAGGTGGCTTCTATGCTCAGGAAACTTGGGGTACCGCCTTGGAGTCTGGAGACAT 7488
Db 7441 AATAGGTGGCTTCTATGCTCAGGAAACTTGGGGTACCGCCTTGGAGTCTGGAGACAT 7500
QY 7489 CGGGCCAGAAAGTGTCCGGCTAGGCTACTGTCCAGAGGGGGAGGGCTGCCACTTGTGGC 7548
Db 7501 CGGGCCAGAAAGTGTCCGGCTAGGCTACTGTCCAGAGGGGGAGGGCTGCCACTTGTGGC 7560
QY 7549 AAGTACCTCTTCAACTGGGCAAGAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7608
Db 7561 AAGTACCTCTTCAACTGGGCAAGAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7620
QY 7609 TCCAGTTTGGATTTATCCAGCTGGTTCGTGTGCTGCTTACAGCGGGGAGACATATAC 7668
Db 7621 TCCAGTTTGGATTTATCCAGCTGGTTCGTGTGCTTACAGCGGGGAGACATATAC 7680
QY 7669 AGCCTGTCTGTCGCCGACCCCGCTGGTTCACTGTGTGCTACTCTCTCTCTGTAGGG 7728
Db 7681 AGCCTGTCTGTCGCCGACCCCGCTGGTTCACTGTGTGCTACTCTCTCTCTGTAGGG 7740
QY 7729 GTAGGATCTATCTACTTCCCAACCGATGAACGGGACCTAAACACTCCAGGCAATAGG 7788
Db 7741 GTAGGATCTATCTACTTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCAATAGG 7800
QY 7789 CCATCCTGTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7848
Db 7801 CCATCCTGTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860
QY 7849 TTTTCTCCTTTTTTTTCTCTTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7908
Db 7861 TTTTCTCCTTTTTTTTCTCTTTTTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
QY 7909 CCTAGTCAGGCTAGCTGTGAAGGTCCGTGAGCCGCTGACTGCGAGAGTCTGATAC 7968
Db 7921 CCTAGTCAGGCTAGCTGTGAAGGTCCGTGAGCCGCTGACTGCGAGAGTCTGATAC 7980
QY 7969 TGGCCTCTCTGCAGATCAAGT 7989
Db 7981 TGGCCTCTCTGCAGATCAAGT 8001

RESULT 11
AX036257
LOCUS AX036257 8001 bp DNA linear PAT 16-NOV-2000
DEFINITION Sequence 6 from Patent EP1043399.
ACCESSION AX036257
VERSION AX036257.1 GI:11225873
KEYWORDS Hepatitis C virus
SOURCE Hepatitis C virus
ORGANISM Hepatitis C virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;
Hepacivirus.
REFERENCE 1
AUTHORS Bartenschlager,R.D.
TITLE Hepatitis c virus cell culture system
JOURNAL Patent: EP 1043399-A 6 11-OCT-2000;
BARTENSCHLAGER RALF DR (DE)
FEATURES
source
1. .8001
/organism="Hepatitis C virus"
/mol_type="unassigned DNA"
/db_xref="taxon:11103"

ORIGIN
Query Match 99.5%; Score 7946.2; DB 6; Length 8001;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 7976; Conservative 0; Mismatches 13; Indels 12; Gaps 1;
QY 1 GCCAGCCCCCGATTGGGGCGGACATCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
Db 1 GCCAGCCCCCGATTGGGGCGGACATCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
QY 61 TCTTCAAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTCGACGCTCCAGGAC 120
```

Db 1141 CTCCGATTTCGACGGCATCGCCTTCTATCGCCTTCTTGACGAGTTCTTCTGAGTTTAAA 1200  
QY 1189 CAGACCAACACGGTTTCCCTCTAGCGGATCAATTCGCCCCCTCTCCCTCCGCCCCCT 1248  
Db 1201 CAGACCAACACGGTTTCCCTCTAGCGGATCAATTCGCCCCCTCTCCCTCCGCCCCCT 1260  
QY 1249 AACGTTACTGCGCAAGCGCTTGGAAATAAGCGCGGTGTGGTTTGTCTATATGTTATT 1308  
Db 1261 AACGTTACTGCGCAAGCGCTTGGAAATAAGCGCGGTGTGGTTTGTCTATATGTTATT 1320  
QY 1309 TCCACCAATATTCGCGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGGCCCTGTCTCTTG 1368  
Db 1321 TCCACCAATATTCGCGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGGCCCTGTCTCTTG 1380  
QY 1369 ACAGGATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGTAATGTC 1428  
Db 1381 ACAGGATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGTAATGTC 1440  
QY 1429 GTGAAGGAAGCAGTTCCCTCTGGAAGCTTCTTGAAGACAAACAACTGTAGCGACCCCT 1488  
Db 1441 GTGAAGGAAGCAGTTCCCTCTGGAAGCTTCTTGAAGACAAACAACTGTAGCGACCCCT 1500  
QY 1489 TGCAGGACGCGAAACCCCTCGGACAGGTGCTCTCGGCCAAAGCCACGTGTA 1548  
Db 1501 TGCAGGACGCGAAACCCCTCGGACAGGTGCTCTCGGCCAAAGCCACGTGTA 1560  
QY 1549 TAAATACACCTGCAAGGGGCAACCCCAAGTGCACCGTGTGAGTTGATGTTGTG 1608  
Db 1561 TAAATACACCTGCAAGGGGCAACCCCAAGTGCACCGTGTGAGTTGATGTTGTG 1620  
QY 1609 GAAAGGTCAAAATGGCTCTCTCAAGCGTATTCACAGGGGCTGAGGATGCCAGAG 1668  
Db 1621 GAAAGGTCAAAATGGCTCTCTCAAGCGTATTCACAGGGGCTGAGGATGCCAGAG 1680  
QY 1669 GTACCCCATTTGATGGGATCTGATCTGGGGCTCGGTGCAATGTTTACATGTTTGTAG 1728  
Db 1681 GTACCCCATTTGATGGGATCTGATCTGGGGCTCGGTGCAATGTTTACATGTTTGTAG 1740  
QY 1729 TCGAGGTTAAAAACGTCCTAGGCCCGCCGAAACCAAGGGGAGCTGGTTTCTTTGAAAA 1788  
Db 1741 TCGAGGTTAAAAACGTCCTAGGCCCGCCGAAACCAAGGGGAGCTGGTTTCTTTGAAAA 1800  
QY 1789 CAGGATAATACCATGGCGCTTATTAAGCGCTTACTCCCAACAGACGCGAGGCTACTTGGC 1848  
Db 1801 CAGGATAATACCATGGCGCTTATTAAGCGCTTACTCCCAACAGACGCGAGGCTACTTGGC 1860  
QY 1849 TGCATCATCTACTAGCCTCACAGGCGGGACAGGAACCAAGTCTGAGGGGAGGTCCAAAGTG 1908  
Db 1861 TGCATCATCTACTAGCCTCACAGGCGGGACAGGAACCAAGTCTGAGGGGAGGTCCAAAGTG 1920  
QY 1909 GTCTCCACCGCAACACAAATCTTTCTGGGACCTTGGCTCAATGGCGTGTGTTGACGTGTC 1968  
Db 1921 GTCTCCACCGCAACACAAATCTTTCTGGGACCTTGGCTCAATGGCGTGTGTTGACGTGTC 1980  
QY 1969 TATCATGTGCGGCTCAAAAGACCTTTGCGGCCCAAGGGCCCAATCACCCAAATGTAC 2028  
Db 1981 TATCATGTGCGGCTCAAAAGACCTTTGCGGCCCAAGGGCCCAATCACCCAAATGTAC 2040  
QY 2029 ACCAATGTGAACAGGACCTCGTGGCTGGCAAGCGCCCGCGGGGCGGTCTCTTGACA 2088  
Db 2041 ACCAATGTGAACAGGACCTCGTGGCTGGCAAGCGCCCGCGGGGCGGTCTCTTGACA 2100  
QY 2089 CCATGACCTGCGGACGCTCGGACCTTTACTTGGTACGAGGATGCGGATGTCATTCGG 2148  
Db 2101 CCATGACCTGCGGACGCTCGGACCTTTACTTGGTACGAGGATGCGGATGTCATTCGG 2160  
QY 2149 GTGCGCGGGGCGGACAGCAGGGGAGCTACTCTCCCGAGGCGCTCTCTACTTGT 2208  
Db 2161 GTGCGCGGGGCGGACAGCAGGGGAGCTACTCTCCCGAGGCGCTCTCTACTTGT 2220  
QY 2209 AAGGCTCTTTCGGGCGGTCCACTGCTCTGCCCTTCGGGGCAGCGTGTGGGCACTTTTCG 2268  
Db 2221 AAGGCTCTTTCGGGCGGTCCACTGCTCTGCCCTTCGGGGCAGCGTGTGGGCACTTTTCG 2280

Db 61 TCTTACCGCAAGACGCTAGCCATAGGTTAGTATAGTGTGTCGAGCTCCAGGAC 120  
QY 121 CCCCTCCCGGAGAGCATAAGTGTCTGGGAACCGGTGAGTACACCGGAATTCGCGAG 180  
Db 121 CCCCTCCCGGAGAGCATAAGTGTCTGGGAACCGGTGAGTACACCGGAATTCGCGAG 180  
QY 181 GACGACCGGTCCTTTCTTGGATCAACCGCTCAATGCTCGGAGATTTGGGGTGCCTCC 240  
Db 181 GACGACCGGTCCTTTCTTGGATCAACCGCTCAATGCTCGGAGATTTGGGGTGCCTCC 240  
QY 241 CGGAGACTGTAGCGAGTGTGTTGGTTCGGAAGGCTTGTGGTACTGCTGATAGG 300  
Db 241 CGGAGACTGTAGCGAGTGTGTTGGTTCGGAAGGCTTGTGGTACTGCTGATAGG 300  
QY 301 GTGCTTCGAGTGCCTCGGAGGCTCTGTAGACCGTGCACATGAGCAGAAATCTCAAC 360  
Db 301 GTGCTTCGAGTGCCTCGGAGGCTCTGTAGACCGTGCACATGAGCAGAAATCTCAAC 360  
QY 361 CTAAAGAAAAACAAA-----GGGCGGCCATGATTGAACAAGATGGATTGC 408  
Db 361 CTAAAGAAAAACAAACGTAACCAACGGCGGCCCATGATTGAACAAGATGGATTGC 420  
QY 409 ACAGAGTTCTCCGCGCTTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAAAGA 468  
Db 421 ACAGAGTTCTCCGCGCTTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAAAGA 480  
QY 469 CAATCGGCTGTCTGATGCGCGCTGTTCGGCTGTGACGCGAGGGCGCGCTTCTTT 528  
Db 481 CAATCGGCTGTCTGATGCGCGCTGTTCGGCTGTGACGCGAGGGCGCGCTTCTTT 540  
QY 529 TTGTCAAGACGACCTGTCGCGTGCCTGAATGAATGCAAGTGCAGGACGAGCAGCGGCTAT 588  
Db 541 TTGTCAAGACGACCTGTCGCGTGCCTGAATGAATGCAAGTGCAGGACGAGCAGCGGCTAT 600  
QY 589 CGTGGCTGGCCACGACGGCGCTTCTTGGCAGCTGTGCTCGAAGTGTCTACTGAAAGCG 648  
Db 601 CGTGGCTGGCCACGACGGCGCTTCTTGGCAGCTGTGCTCGAAGTGTCTACTGAAAGCG 660  
QY 649 GAAAGGACTGCTGCTTATGGGCGAAGTGCAGGCGAGGATCTCTGTGATCTCACTTGTG 708  
Db 661 GAAAGGACTGCTGCTTATGGGCGAAGTGCAGGCGAGGATCTCTGTGATCTCACTTGTG 720  
QY 709 CTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTGATC 768  
Db 721 CTCTGCGGAGAAAGTATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTGATC 780  
QY 769 CGGCTACTGCGCATTCGACCAACGAGAAACATCGCATCGAGCGAGCAGTACTCGGA 828  
Db 781 CGGCTACTGCGCATTCGACCAACGAGAAACATCGCATCGAGCGAGCAGTACTCGGA 840  
QY 829 TGGAGCGGCTGCTGATCAGATGATCTGAGCAGAGATCTGAGGCGTCCGCGCAG 888  
Db 841 TGGAGCGGCTGCTGATCAGATGATCTGAGCAGAGATCTGAGGCGTCCGCGCAG 900  
QY 889 CCGAATCTGTCGCGAGGCTCAAGCGCGCATGCCCCGAGGAGATCTGCTGCTGACCC 948  
Db 901 CCGAATCTGTCGCGAGGCTCAAGCGCGCATGCCCCGAGGAGATCTGCTGCTGACCC 960  
QY 949 ATGCGGATGCTGCTTGGCGAATATGTTGGGAAATGCGCGCTTTTCTGGATTCATCG 1008  
Db 961 ATGCGGATGCTGCTTGGCGAATATGTTGGGAAATGCGCGCTTTTCTGGATTCATCG 1020  
QY 1009 ACTGTGCGGCTGCTGAGCGGACCGCTATCAGGACATAGCGTTGCTACCGTGATA 1068  
Db 1021 ACTGTGCGGCTGCTGAGCGGACCGCTATCAGGACATAGCGTTGCTACCGTGATA 1080  
QY 1069 TTGTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTGCTTTTACGGTATCGCG 1128  
Db 1081 TTGTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTGCTTTTACGGTATCGCG 1140  
QY 1129 CTCCCGATTCGACGCGATCGCCTTCTATCGCCTTCTGAGAGTCTTCTGAGTTTAAA 1188

QY 2269 GCTCGCTGTGACCCGAGGGGTTGCGAAGGCGGTGGAATTGTATCCCGTCGAGTCTATG 2328  
DB 2281 GCTGCGGTGTGCACCGAGGGGTTGCGAAGGCGGTGGAATTGTATCCCGTCGAGTCTATG 2340  
QY 2329 GAAACCACTATGCGGTCCCGGTCTTTCAGGACAACTCGTCCCTCCCGCGGTACCGGAG 2388  
DB 2341 GAAACCACTATGCGGTCCCGGTCTTTCAGGACAACTCGTCCCTCCCGCGGTACCGGAG 2400  
QY 2389 ACATATCCAGGTGCCCATCTACACGCCCTCTAGGTAGCGGCAAGAGACACTAAGGTGCGG 2448  
DB 2401 ACATATCCAGGTGCCCATCTACACGCCCTCTAGGTAGCGGCAAGAGACACTAAGGTGCGG 2460  
QY 2449 GCTGGTATGACGCCCAAGGATTAAGGTGCTTGTCTGAAACCGGTGCGTCCCGGCGACC 2508  
DB 2461 GCTGGTATGACGCCCAAGGATTAAGGTGCTTGTCTGAAACCGGTGCGTCCCGGCGACC 2520  
QY 2509 CTAGGTTTGGGGCGGTATATGTCTAAGGCACATGGTATCGACCTTAAACATCAGAACCGGG 2568  
DB 2521 CTAGGTTTGGGGCGGTATATGTCTAAGGCACATGGTATCGACCTTAAACATCAGAACCGGG 2580  
QY 2569 GTAAGGACCATCACACGGGTGCCCGCATACGTACTCCACCTATGGCAAGTTTCTTGCC 2628  
DB 2581 GTAAGGACCATCACACGGGTGCCCGCATACGTACTCCACCTATGGCAAGTTTCTTGCC 2640  
QY 2629 GACGGTGGTGTCTGCGGGGCGCTATGACATCAATAATGATGATGACCACTCAACT 2688  
DB 2641 GACGGTGGTGTCTGCGGGGCGCTATGACATCAATAATGATGATGACCACTCAACT 2700  
QY 2689 GACTCGACCACTATCTCTGGGCACTCGGACAGTCTGGACCAAGCGGAGACGGCTGGAGG 2748  
DB 2701 GACTCGACCACTATCTCTGGGCACTCGGACAGTCTGGACCAAGCGGAGACGGCTGGAGG 2760  
QY 2749 CGACTCGTGTCTGCCACCGCTACGCTCCGGATCGGTCAACGTGCCACATCCAAC 2808  
DB 2761 CGACTCGTGTCTGCCACCGCTACGCTCCGGATCGGTCAACGTGCCACATCCAAC 2820  
QY 2809 ATCGAGGAGTGGTCTGTGACGACTGGAGAAATCCCTTTATGGCAAGCCATCCCC 2868  
DB 2821 ATCGAGGAGTGGTCTGTGACGACTGGAGAAATCCCTTTATGGCAAGCCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGAGGACCTCATTTCTGCCATTCGAAGAAAGTGTAT 2928  
DB 2881 ATCGGCGACCATCAAGGGGGAGGACCTCATTTCTGCCATTCGAAGAAAGTGTAT 2940  
QY 2929 GAGCTCGCGCGAAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGCGCTT 2988  
DB 2941 GAGCTCGCGCGAAGCTATCCGGCTCGGACTCAATGTGTAGCATATTAACGGGCGCTT 3000  
QY 2989 GATGTATCGGTATACCAACTAGCGGAGACGTCAATGTGTAGCAACGGACGCTCTAATG 3048  
DB 3001 GATGTATCGGTATACCAACTAGCGGAGACGTCAATGTGTAGCAACGGACGCTCTAATG 3060  
QY 3049 ACGGCTTTACCGCGATTTGACTCAGTGATCGACTGCAATACATGTCTACCCAGACA 3108  
DB 3061 ACGGCTTTACCGCGATTTGACTCAGTGATCGACTGCAATACATGTCTACCCAGACA 3120  
QY 3109 ETCGACTTCAGCTTGACCCGACCTTCACTATGACGACGACCGGTGCCAAGACGG 3168  
DB 3121 ETCGACTTCAGCTTGACCCGACCTTCACTATGACGACGACCGGTGCCAAGACGG 3180  
QY 3169 GTGTACGCTCGAGCGGCGAGGAGACTGGTAGGGGAGGATGGGCATTTACAGGTTT 3228  
DB 3181 GTGTACGCTCGAGCGGCGAGGAGACTGGTAGGGGAGGATGGGCATTTACAGGTTT 3240  
QY 3229 GTGACTCCAGGAGAACGGCCCTCGGSCATGTTGATTCTCGGTTCTGTGCGAGTGCTAT 3288  
DB 3241 GTGACTCCAGGAGAACGGCCCTCGGSCATGTTGATTCTCGGTTCTGTGCGAGTGCTAT 3300  
QY 3289 GACGCGGGCTGTGCTTGTGTACGAGCTCACGCCCGCGAGACCTCAGTTAGGTTGCGGGCT 3348  
DB 3301 GACGCGGGCTGTGCTTGTGTGTACGAGCTCACGCCCGCGAGACCTCAGTTAGGTTGCGGGCT 3360

QY 3349 TACCTAAACACACGAGGGTTGCCGTCTGCCAGGACCACTCTGGAGTTCTGGGAGAGCGTC 3408  
DB 3361 TACCTAAACACACGAGGGTTGCCGTCTGCCAGGACCACTCTGGAGTTCTGGGAGAGCGTC 3420  
QY 3409 TTTACAGGCTTCAACCATAGACCCCATTTCTTGTCCAGACTAAGCAGGACAGAGAC 3468  
DB 3421 TTTACAGGCTTCAACCATAGACCCCATTTCTTGTCCAGACTAAGCAGGACAGAGAC 3480  
QY 3469 AACTTCCCTACCTGGTAGCATACAGGCTACGGTGTGGCCAGGGCTCAGGCTCAACT 3528  
DB 3481 AACTTCCCTACCTGGTAGCATACAGGCTACGGTGTGGCCAGGGCTCAGGCTCAACT 3540  
QY 3529 CCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAGACCTACGCTGCAACGGCCA 3588  
DB 3541 CCATCGTGGGACCAATGTGGAAGTGTCTCATACGGCTAAGACCTACGCTGCAACGGCCA 3600  
QY 3589 ACGCCCTTGTGTATAGGCTGGGAGCCCTCAAAACAGAGTTACTACCAACACACCCCATATA 3648  
DB 3601 ACGCCCTTGTGTATAGGCTGGGAGCCCTCAAAACAGAGTTACTACCAACACACCCCATATA 3660  
QY 3649 ACCAAATACATATGGCATGATGCGCTGACCTGGAGGTGTCACGAGACCTGGGTG 3708  
DB 3661 ACCAAATACATATGGCATGATGCGCTGACCTGGAGGTGTCACGAGACCTGGGTG 3720  
QY 3709 CTGGTAGGGGAGTCTTAGCAGCTCTGGCCGCTATTGCTGACACAGGACGCTGCTC 3768  
DB 3721 CTGGTAGGGGAGTCTTAGCAGCTCTGGCCGCTATTGCTGACACAGGACGCTGCTC 3780  
QY 3769 ATTGTGGGCGAGGATCATCTTGTCCGGAAGCGCGGCATCATTCGCCGACAGGGAAGTCTT 3828  
DB 3781 ATTGTGGGCGAGGATCATCTTGTCCGGAAGCGCGGCATCATTCGCCGACAGGGAAGTCTT 3840  
QY 3829 TACCGGAGTTCGATGAGATGGAAGTGGCTCACACCTTCCCTTACATCGAAACAGGA 3888  
DB 3841 TACCGGAGTTCGATGAGATGGAAGTGGCTCACACCTTCCCTTACATCGAAACAGGA 3900  
QY 3889 ATGCGAGCTCGCGGACCAATTCAAACAGAGGCAATCGGTTGCTGCAACACGACCCACAG 3948  
DB 3901 ATGCGAGCTCGCGGACCAATTCAAACAGAGGCAATCGGTTGCTGCAACACGACCCACAG 3960  
QY 3949 CAAGCGGAGGCTGTGCTCCCGTGGTGAATCCAAAGTGGCGGACCTCGAAGCCTTCTGG 4008  
DB 3961 CAAGCGGAGGCTGTGCTCCCGTGGTGAATCCAAAGTGGCGGACCTCGAAGCCTTCTGG 4020  
QY 4009 GCGAAGCATATGTGGAATTTATCAGCGGATACAAATATTTAGCAGGCTGTGCCACTGTG 4068  
DB 4021 GCGAAGCATATGTGGAATTTATCAGCGGATACAAATATTTAGCAGGCTGTGCCACTGTG 4080  
QY 4069 CCTGGCAACCCGCGATAGCATCACTGATGGGATTCACAGCCTCTATCACAGCCGCTC 4128  
DB 4081 CCTGGCAACCCGCGATAGCATCACTGATGGGATTCACAGCCTCTATCACAGCCGCTC 4140  
QY 4129 ACCACCCAAACATACCCTCTGTTTAAACATCTCTGGGGGATGGGTGGCCGCCAACTTGCT 4188  
DB 4141 ACCACCCAAACATACCCTCTGTTTAAACATCTCTGGGGGATGGGTGGCCGCCAACTTGCT 4200  
QY 4189 CTTCCAGCGCTGTCTGTCTTTCGTAGCGCCGCACTGCTGGAGCGGCTGTGGCAGC 4248  
DB 4201 CTTCCAGCGCTGTCTGTCTTTCGTAGCGCCGCACTGCTGGAGCGGCTGTGGCAGC 4260  
QY 4249 ATAGGCTTGGGAAGGCTTGTGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGC 4308  
DB 4261 ATAGGCTTGGGAAGGCTTGTGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGC 4320  
QY 4309 GGGCTCTGGCCCTTTAAGGTATAGCGCGGAGATGCCCTCCACCGAGACCTGGTTAAC 4368  
DB 4321 GGGCTCTGGCCCTTTAAGGTATAGCGCGGAGATGCCCTCCACCGAGACCTGGTTAAC 4380  
QY 4369 CTACTCCCTGTCTATCCTCTCCCTGGCGCCTTAGTCTGTCGGGGTCTGTGTCGACGATA 4428  
DB 4381 CTACTCCCTGTCTATCCTCTCCCTGGCGCCTTAGTCTGTCGGGGTCTGTGTCGACGATA 4440  
QY 4429 CTGCGTGGCAGCTGGGCCCGGAGGGGCTGTGTCAGTGATGAACCGGCTGATAGC 4488





```
QY 6649 TGTTTTGAATCAACGGTCACTGAGAAATGACATCCGTGTTGAGGAGTCAATCTACCAATGT 6708
Db 6661 TGTTTTGAATCAACGGTCACTGAGAAATGACATCCGTGTTGAGGAGTCAATCTACCAATGT 6720
QY 6709 TGTGACTTGGCCCGGAGAGCCAGACAGGCAATAAGTCCGTCACAGAGCGGCTTTACATC 6768
Db 6721 TGTGACTTGGCCCGGAGAGCCAGACAGGCAATAAGTCCGTCACAGAGCGGCTTTACATC 6780
QY 6769 GGGGGCCCGCTGACTAAATTTAAAGGAGAACTCGGCTATCGCCGCTATCGCCGCGAGC 6828
Db 6781 GGGGGCCCGCTGACTAAATTTAAAGGAGAACTCGGCTATCGCCGCTATCGCCGCGAGC 6840
QY 6829 GGTGTACTGACGACGACGCTGCGGTAAATACCTCTACATGTTACTTGAAGCCGCTCGGCGC 6888
Db 6841 GGTGTACTGACGACGACGCTGCGGTAAATACCTCTACATGTTACTTGAAGCCGCTCGGCGC 6900
QY 6889 TGTCCAGCTGCGAAGCTCCAGGACTGCACGATGCTCGTATGGGAGACGACCTTCGTGT 6948
Db 6901 TGTCCAGCTGCGAAGCTCCAGGACTGCACGATGCTCGTATGGGAGACGACCTTCGTGT 6960
QY 6949 ATCTGTAAGGCGGGGACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7008
Db 6961 ATCTGTAAGGCGGGGACCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7020
QY 7009 ATGACTAGACTCTGCCCCCTTGGGAGCCGCGCCAAACCAAGAAATACGACTTGGAGTTG 7068
Db 7021 ATGACTAGACTCTGCCCCCTTGGGAGCCGCGCCAAACCAAGAAATACGACTTGGAGTTG 7080
QY 7069 ATAAATCATGCTCCCAATGTGTGAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7128
Db 7081 ATAAATCATGCTCCCAATGTGTGAGTGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7140
QY 7129 TATCTCACCGGTGACCCACCAACCCCTTGGCGGGGCTGCGTGGGAGACAGCTAGACAC 7188
Db 7141 TATCTCACCGGTGACCCACCAACCCCTTGGCGGGGCTGCGTGGGAGACAGCTAGACAC 7200
QY 7189 ACTCAGTCAATCTCTGGCTAGGAAACATCATCATGTATGCGCCCACTTGTGGGCAAGG 7248
Db 7201 ACTCAGTCAATCTCTGGCTAGGAAACATCATCATGTATGCGCCCACTTGTGGGCAAGG 7260
QY 7249 ATGATCTGTAGTCACTTCTTCTCCATCTTCTAGCTCAGGAAACAACTTCAAAAGCC 7308
Db 7261 ATGATCTGTAGTCACTTCTTCTCCATCTTCTAGCTCAGGAAACAACTTCAAAAGCC 7320
QY 7309 CTAGATTTGTAGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7368
Db 7321 CTAGATTTGTAGTCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 7380
QY 7369 ATTCAAGACTCCATGGGCTTACGGGATTTTCACTCCATAGTTACTCTCAGGTGAGATC 7428
Db 7381 ATTCAAGACTCCATGGGCTTACGGGATTTTCACTCCATAGTTACTCTCAGGTGAGATC 7440
QY 7429 AATAGGTGGCTTCATGCTCAGGAAACTTGGGTAACCCGCTTGGAGTCTGGAGACAT 7488
Db 7441 AATAGGTGGCTTCATGCTCAGGAAACTTGGGTAACCCGCTTGGAGTCTGGAGACAT 7500
QY 7489 CGGGCCAGAAAGTGTCCGCTAGGCTTACTGTGCCAGGGGGAGGAGGAGGAGGAGGAGGAGG 7548
Db 7501 CGGGCCAGAAAGTGTCCGCTAGGCTTACTGTGCCAGGGGGAGGAGGAGGAGGAGGAGGAGG 7560
QY 7549 AAGTACCTCTCAATCGGAGTAAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGGG 7608
Db 7561 AAGTACCTCTCAATCGGAGTAAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGGG 7620
QY 7609 TCCAGTTCGATTAATCCAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 7668
Db 7621 TCCAGTTCGATTAATCCAGCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 7680
QY 7669 AGCCTGTCTCGTCCCGACCCCGCTGTTTCACTGTGTGTGCTTACTCTTCTTCTGTAGGG 7728
Db 7681 AGCCTGTCTCGTCCCGACCCCGCTGTTTCACTGTGTGTGCTTACTCTTCTTCTGTAGGG 7740
```

```
QY 7729 CTAGCATCTATCTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGG 7788
Db 7741 CTAGCATCTATCTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGG 7800
QY 7789 CCATCCCTCTTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7848
Db 7801 CCACTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
QY 7849 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7908
Db 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
QY 7909 CCTAGTCAAGCTAGCTGTAAGGCTCGTGAAGCTCGTGAAGCTCGTGAAGCTCGTGAAGCT 7968
Db 7921 CCTAGTCAAGCTAGCTGTAAGGCTCGTGAAGCTCGTGAAGCTCGTGAAGCTCGTGAAGCT 7980
QY 7969 TGGCTCTCTCTGAGATCAAGT 7989
Db 7981 TGGCTCTCTCTGAGATCAAGT 8001

RESULT 12
AR406049
LOCUS AR406049 8001 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 28 from patent US 6630343.
ACCESSION AR406049
VERSION AR406049.1 GI:40155176
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 8001)
AUTHORS Bartenschlager,R.
TITLE Hepatitis C virus culture system
JOURNAL Patent: US 6630343-A 28 07-OCT-2003;
FEATURES
source
1. 8001
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 99.4%; Score 7939.8; DB 6; Length 8001;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 7972; Conservative 0; Mismatches 17; Indels 12; Gaps 1;

QY 1 GCCAGCCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
Db 1 GCCAGCCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60

QY 61 TCTTCACGAGAAAGCGTCTAGCCATGGCTTAGTATGAGTGTGTCGAGCCTCCAGGAC 120
Db 61 TCTTCACGAGAAAGCGTCTAGCCATGGCTTAGTATGAGTGTGTCGAGCCTCCAGGAC 120

QY 121 CCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTAGTACACCGGAATTCGAC 180
Db 121 CCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTAGTACACCGGAATTCGAC 180

QY 181 GACGACCGGGTCTTTTCTTTGGATCAACCGCTCAATGCTGGAGATTTGGGCGTCCCC 240
Db 181 GACGACCGGGTCTTTTCTTTGGATCAACCGCTCAATGCTGGAGATTTGGGCGTCCCC 240

QY 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGGCTTGTGTTACTGCTGATAGG 300
Db 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGGCTTGTGTTACTGCTGATAGG 300

QY 301 GTGCTTGGAGTGTGCGGAGGCTCTGTAGACCGTGCACCATGAGCAGATTCCTAAC 360
Db 301 GTGCTTGGAGTGTGCGGAGGCTCTGTAGACCGTGCACCATGAGCAGATTCCTAAC 360

QY 361 CTCAGAGAAACCAACCAACGTAACCAACGCGGCGGCGCCATGATTGAAACAGATGATTCG 408
Db 361 CTCAGAGAAACCAACCAACGTAACCAACGCGGCGGCGCCATGATTGAAACAGATGATTCG 420
```

QY 409 ACGAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGTATGACTGGGCACAACAGA 468  
Db 421 ACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATTCCGGTATGACTGGGCACAACAGA 480  
QY 469 CAATCGGCTGCTCTGATGCCCGCGTGTCCGGCTGTACGCGCAGGGGCGCCGGTCTTTT 528  
Db 481 CAATCGGCTGCTCTGATGCCCGCGTGTCCGGCTGTACGCGCAGGGGCGCCGGTCTTTT 540  
QY 529 TTGTCAGACGACCTGTCCGGTCCCTGATGAATGAATGACGACGAGGACGCGGCTAT 588  
Db 541 TTGTCAGACGACCTGTCCGGTCCCTGATGAATGAATGACGACGAGGACGCGGCTAT 600  
QY 589 CGTGGCTGGCCACAGACGGGCGTTCCTTCCGAGCTGTCTCGACGTTGTCTACTGAAGGG 648  
Db 601 CGTGGCTGGCCACAGACGGGCGTTCCTTCCGAGCTGTCTCGACGTTGTCTACTGAAGGG 660  
QY 649 GAAAGGACTGCGTGTATTTGGGCGAAGTTCGCGGGCAGGATCTCCTGTCATCTCACCTTG 708  
Db 661 GAAAGGACTGCGTGTATTTGGGCGAAGTTCGCGGGCAGGATCTCCTGTCATCTCACCTTG 720  
QY 709 CTCCTGCCGAAAGATATCCATCATGCTGATGCAATGCGGGCGCTGCATACGCTTGATC 768  
Db 721 CTCCTGCCGAAAGATATCCATCATGCTGATGCAATGCGGGCGCTGCATACGCTTGATC 780  
QY 769 CGGCTACCTGCCCATTTCCACCAACGAAACATCGCATCGAGCGACGTAATCTCGGA 828  
Db 781 CGGCTACCTGCCCATTTCCACCAACGAAACATCGCATCGAGCGACGTAATCTCGGA 840  
QY 829 TGGAGCGGCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAG 888  
Db 841 TGGAGCGGCTTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAG 900  
QY 889 CCGAAGCTTTCGCGAGCTCAAGCGCGCATGCGCGGAGGATCTCGTCGACCC 948  
Db 901 CCGAAGCTTTCGCGAGCTCAAGCGCGCATGCGCGGAGGATCTCGTCGACCC 960  
QY 949 ATGCGGATGCTGCTGCTGCCAATATCATGTGGAAATGCGCGCTTTTCTGGATTATCG 1008  
Db 961 ATGCGGATGCTGCTTGCAGATATCATGTGGAAATGCGCGCTTTTCTGGATTATCG 1020  
QY 1009 ACTGTGCGGCTGGGTGTGGCGAACCGCTATACGACATAGCGCTTGGCTACCGGTGATA 1068  
Db 1021 ACTGTGCGGCTGGGTGTGGCGAACCGCTATACGACATAGCGCTTGGCTACCGGTGATA 1080  
QY 1069 TTGCTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCG 1128  
Db 1081 TTGCTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCG 1140  
QY 1129 CTCGCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAA 1188  
Db 1141 CTCGCGATTGCGAGCGCATCGCCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAA 1200  
QY 1189 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCGCTCTCCTCCCCCCCCCT 1248  
Db 1201 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCGCTCTCCTCCCCCCCCCT 1260  
QY 1249 AACCTTACTGGCGAAGCCCTTGGATTAAGCCGCTGTGGTTGTGCTATGCTATTTT 1308  
Db 1261 AACCTTACTGGCGAAGCCCTTGGATTAAGCCGCTGTGGTTGTGCTATGCTATTTT 1320  
QY 1309 TCACACCATATGCGGCTTTTGGCAATGTGAGGCGCCGGAACCTGCGCCCTGTCTTCTTG 1368  
Db 1321 TCACACCATATGCGGCTTTTGGCAATGTGAGGCGCCGGAACCTGCGCCCTGTCTTCTTG 1380  
QY 1369 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCCAAAGGATGCAAGTCTGTGGAATGTC 1428  
Db 1381 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCCAAAGGATGCAAGTCTGTGGAATGTC 1440  
QY 1429 GTGAAGGAAGCAGTTCTCTGGAAGCTTCTTTGAAGACAAACAACTGTGAGCGACCTT 1488  
Db 1441 GTGAAGGAAGCAGTTCTCTGGAAGCTTCTTTGAAGACAAACAACTGTGAGCGACCTT 1500  
QY 1489 TGCAGGACGGAACCCCCACCTGCGGACAGGTGCTCTGCGGCCAAAAAGCCACGTGTA 1548

Db 1501 TGCAGGACGGAACCCCCACCTTGGGACAGGTGCTCTGCGGCCAAAAAGCCACGTGTA 1560  
QY 1549 TAAGATACACTGCAAAAGGGGCAAAACCCAGTGCACCTGTGTAGTTGATAGTTGTG 1608  
Db 1561 TAAGATACACTGCAAAAGGGGCAAAACCCAGTGCACCTGTGTAGTTGATAGTTGTG 1620  
QY 1609 GAAAGAGTCAAAATGGCTCTCCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAGAAG 1668  
Db 1621 GAAAGAGTCAAAATGGCTCTCCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAGAAG 1680  
QY 1669 GTACCCCATTTATGGGATCTGATCTGGGCTCGGTGCACATCTTTACATGTGTTTAG 1728  
Db 1681 GTACCCCATTTATGGGATCTGATCTGGGCTCGGTGCACATCTTTACATGTGTTTAG 1740  
QY 1729 TCGAGGTTAAAAAACGCTTAGGCCCCCGCAACACCGGGGACGTGTTTCTTTGAAAAA 1788  
Db 1741 TCGAGGTTAAAAAACGCTTAGGCCCCCGCAACACCGGGGACGTGTTTCTTTGAAAAA 1800  
QY 1789 CACGATTAATACATGGCGCTTATTACGGCTTACTCCCAACAGACGCGAGGCTACTTGGC 1848  
Db 1801 CACGATTAATACATGGCGCTTATTACGGCTTACTCCCAACAGACGCGAGGCTACTTGGC 1860  
QY 1849 TGCATCATCTAGCCTCAAGGCGGACAGGAACACAGGTTCGAGGGGAGGTCCAAGTG 1908  
Db 1861 TGCATCATCTAGCCTCAAGGCGGACAGGAACACAGGTTCGAGGGGAGGTCCAAGTG 1920  
QY 1909 GTCTCCACCGCAACACAATCTTCTCGCGACCTGCGTCAATGCGCTGTGTGACTGTC 1968  
Db 1921 GTCTCCACCGCAACACAATCTTCTCGCGACCTGCGTCAATGCGCTGTGTGACTGTC 1980  
QY 1969 TATCATGTCGCGCTCAAGAGCCCTTTCGCGGCCCAAGGGCCCAATCACCAATGTAC 2028  
Db 1981 TATCATGTCGCGCTCAAGAGCCCTTTCGCGGCCCAAGGGCCCAATCACCAATGTAC 2040  
QY 2029 ACCAATGTGACACGAGACCTCTGCGCTGGCAAGCGCCCCCGGGCGCTTCTTGACA 2088  
Db 2041 ACCAATGTGACACGAGACCTCTGCGCTGGCAAGCGCCCCCGGGCGCTTCTTGACA 2100  
QY 2089 CCATGCACTTGGCGAGCTCGACCTTTTACTTGTGTACAGAGGCATGCGATTCCTCG 2148  
Db 2101 CCATGCACTTGGCGAGCTCGACCTTTTACTTGTGTACAGAGGCATGCGATTCCTCG 2160  
QY 2149 GTGCGCGCGGCGGACAGAGGGGAGCTACTCTCCCGGAGGCGCTTCTCTACTTG 2208  
Db 2161 GTGCGCGCGGCGGACAGAGGGGAGCTACTCTCCCGGAGGCGCTTCTCTACTTG 2220  
QY 2209 AAGGGCTCTTTCGCGCGCTCCACTGCTCTGCCCCCTCGGGGCAAGCTGTGGGCTCTTTTCG 2268  
Db 2221 AAGGGCTCTTTCGCGCGCTCCACTGCTCTGCCCCCTCGGGGCAAGCTGTGGGCTCTTTTCG 2280  
QY 2269 GCTGCGCTGTGCACCCGAGGGTTCGAGGCGGTGACCTTTGTAACCGTCTGAGTCTATG 2328  
Db 2281 GCTGCGCTGTGCACCCGAGGGTTCGAGGCGGTGACCTTTGTAACCGTCTGAGTCTATG 2340  
QY 2329 GAAACCACTATCGGCTCCCGGCTTTTACAGGCAACTCGTCCCTCGCGCGGTACCGAG 2388  
Db 2341 GAAACCACTATCGGCTCCCGGCTTTTACAGGCAACTCGTCCCTCGCGCGGTACCGAG 2400  
QY 2389 ACATTCAGGTGGCCCATCTACACGCCCTTACTGTTAGCGGCAAGAGCACTAAGGTGCGG 2448  
Db 2401 ACATTCAGGTGGCCCATCTACACGCCCTTACTGTTAGCGGCAAGAGCACTAAGGTGCGG 2460  
QY 2449 GCTGCGTATGACGCCGAGGGTATAGGTGCTTGTCTCTGAACCGCTCGTCCGCGCGCAC 2508  
Db 2461 GCTGCGTATGACGCCGAGGGTATAGGTGCTTGTCTCTGAACCGCTCGTCCGCGCGCAC 2520  
QY 2509 CTAGGTTTCGGCGTATATGCTTAAGGCATGTTATCGACCTTAACATCAGAAACCGG 2568  
Db 2521 CTAGGTTTCGGCGTATATGCTTAAGGCATGTTATCGACCTTAACATCAGAAACCGG 2580  
QY 2569 GTAAGGACCATCAACCGGTGCCCCCATCAAGCTTCTCCACCTATGGAAGTTTCTTGGC 2628

Db 2581 GTAAGACCATCACACAGGGTGGCCCCCATCAAGTACTCACCTATGGCAAGTTTCTTGCC 2640  
QY 2629 GACGFTGTTGCTCTGGGGCGCCTATGACATCATATAATATGTGATGAGTGCCACTCAACT 2688  
Db 2641 GACGFTGTTGCTCTGGGGCGCCTATGACATCATATAATATGTGATGAGTGCCACTCAACT 2700  
QY 2689 GACTCGACCACTATCTGTGGCATCGGCACAGTCTCTGGACCAAGCGAGAGCGGTGAGCG 2748  
Db 2701 GACTCGACCACTATCTGTGGCATCGGCACAGTCTCTGGACCAAGCGAGAGCGGTGAGCG 2760  
QY 2749 CGACTCGTGTGCTGCCACCGCTACGCTCGGGATCGGTACCGTGCACATCAAAAC 2808  
Db 2761 CGACTCGTGTGCTGCCACCGCTACGCTCGGGATCGGTACCGTGCACATCAAAAC 2820  
QY 2809 ATCGAGAGGTGGCTGTCTCGACATCGGAGAAATCCCTTTTATGGCAAGCCATCCCC 2868  
Db 2821 ATCGAGAGGTGGCTGTCTCGACATCGGAGAAATCCCTTTTATGGCAAGCCATCCCC 2880  
QY 2869 ATCGAGACCATCAAGGGGGAGGACCTCATTTCTGCCATTCRAAGAGAAATGTGAT 2928  
Db 2881 ATCGAGACCATCAAGGGGGAGGACCTCATTTCTGCCATTCRAAGAGAAATGTGAT 2940  
QY 2929 GAGCTCGCGCGAAGCTGTCCGGCCTCGGACTCAATGCTGTAGCATATTAACCGGGCCTT 2988  
Db 2941 GAGCTCGCGCGAAGCTGTCCGGCCTCGGACTCAATGCTGTAGCATATTAACCGGGCCTT 3000  
QY 2989 GATGTATCGGTANATCAACTAGCGGAGAGCTCAATTTGTGTAGCAACGAGCGCTCTAATG 3048  
Db 3001 GATGTATCGGTANATCAACTAGCGGAGAGCTCAATTTGTGTAGCAACGAGCGCTCTAATG 3060  
QY 3049 AGGGCTTTACGGCGATTTGACTCACTGATGATGATGATGATGATGATGATGATGATGATGAT 3108  
Db 3061 AGGGCTTTACGGCGCGACTTGCAGTCACTGATGATGATGATGATGATGATGATGATGATGAT 3120  
QY 3109 GTGCACTTCAGCGCTGACCGGACCTTCACTTTGAGACGACGACCGCTGCCAAGACGCG 3168  
Db 3121 GTGCACTTCAGCGCTGACCGGACCTTCACTTTGAGACGACGACCGCTGCCAAGACGCG 3180  
QY 3169 GTGTACGCTCGAGCGGGAGGACGAGCTGTGTAGGGGAGGATGGGCAATTTACAGGTTT 3228  
Db 3181 GTGTACGCTCGAGCGGGAGGACGAGCTGTGTAGGGGAGGATGGGCAATTTACAGGTTT 3240  
QY 3229 GTGACTTCAGGAGAACGGCCCTCGGCATGTTGATTCCTCGTTCCTGTGCGAGTGCCTAT 3288  
Db 3241 GTGACTTCAGGAGAACGGCCCTCGGCATGTTGATTCCTCGTTCCTGTGCGAGTGCCTAT 3300  
QY 3289 GACGCGGGTGTGCTTGTGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTTGGCGGCT 3348  
Db 3301 GACGCGGGTGTGCTTGTGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTTGGCGGCT 3360  
QY 3349 TACCTAAACACACAGGGTTCGCTGTGCGAGCAATCTGAGTTCTGGGAGAGCGTC 3408  
Db 3361 TACCTAAACACACAGGGTTCGCTGTGCGAGCAATCTGAGTTCTGGGAGAGCGTC 3420  
QY 3409 TTTACAGGCTCACACATAGACGCCCTTCTTCTCCAGACTAAGCAGGCGAGAGAC 3468  
Db 3421 TTTACAGGCTCACACATAGACGCCCTTCTTCTCCAGACTAAGCAGGCGAGAGAC 3480  
QY 3469 AACTTCCCTTACCTGTGTAGCATACAGGCTACCGGTGTGCCAGGCTCAGGCTCCACCT 3528  
Db 3481 AACTTCCCTTACCTGTGTAGCATACAGGCTACCGGTGTGCCAGGCTCAGGCTCCACCT 3540  
QY 3529 CAACTGTGGACCAATGTGGAAAGTGTCTCATACGGCTAAAGCTACGCTGTGACAGGGCCA 3588  
Db 3541 CAACTGTGGACCAATGTGGAAAGTGTCTCATACGGCTAAAGCTACGCTGTGACAGGGCCA 3600  
QY 3589 ACGCCCTGTGTATAGGCTGGAGCGGTTCAAAACGAGTTACTACCAACACACCCCAT 3648  
Db 3601 ACGCCCTGTGTATAGGCTGGAGCGGTTCAAAACGAGTTACTACCAACACACCCCAT 3660  
QY 3649 ACCAAATACATCATGATCATGCTGGCTGACCTGGAGTCTGACGAGCACTGGGTG 3708  
Db 3661 ACCAAATACATCATGATCATGCTGGCTGACCTGGAGTCTGACGAGCACTGGGTG 3720

QY 3709 CTGTAGGCGGAGTCTTAGCAGCTCTGGCGCGGTATTTGCTGTGACAAACAGGACGCTGGTC 3768  
Db 3721 CTGTAGGCGGAGTCTTAGCAGCTCTGGCGCGGTATTTGCTGTGACAAACAGGACGCTGGTC 3780  
QY 3769 ATTGTGGGAGGATCATCTTGTCCGGAAGCCGGCCATCATTTCCCGACAGGGAAGTCCCT 3828  
Db 3781 ATTGTGGGAGGATCATCTTGTCCGGAAGCCGGCCATCATTTCCCGACAGGGAAGTCCCT 3840  
QY 3829 TACCGGAGTTCGATGAGATGGAAGTGGCTCACACCTCCCTTACATCGAACAGGGA 3888  
Db 3841 TACCGGAGTTCGATGAGATGGAAGTGGCTCACACCTCCCTTACATCGAACAGGGA 3900  
QY 3889 ATGAGCTCGCGAAACAATTCAAACAGAAAGCAATCGGGTGTGCAACAGCACCAAG 3948  
Db 3901 ATGAGCTCGCGAAACAATTCAAACAGAAAGCAATCGGGTGTGCAACAGCACCAAG 3960  
QY 3949 CAAGCGAGGCTGCTCTCCCGTGTGGAATCAAGTGGCGGACCTCGAAGCCTTCTGG 4008  
Db 3961 CAAGCGAGGCTGCTCTCCCGTGTGGAATCAAGTGGCGGACCAATCGAAGCCTTCTGG 4020  
QY 4009 GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTG 4068  
Db 4021 GCGAAGCATATGTGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTG 4080  
QY 4069 CTTGGCAACCCCGGATAGATCACTGATGGCAATTCAGAGCTTATACAGCCCGCTC 4128  
Db 4081 CTTGGCAACCCCGGATAGATCACTGATGGCAATTCAGAGCTTATACAGCCCGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTTTTAACTCTGGGGGATGGTGGCCGCCAACTTGT 4188  
Db 4141 ACCACCAACATACCTCTCTTTTAACTCTGGGGGATGGTGGCCGCCAACTTGT 4200  
QY 4189 CTTCCAGCGCTCTCTCTCTTTAGCGCGGATCGCTGGAGCGGTGTGGCAGC 4248  
Db 4201 CTTCCAGCGCTCTCTCTCTTTAGCGCGGATCGCTGGAGCGGTGTGGCAGC 4260  
QY 4249 ATAGGCTTTGGAAAGTGTGTGGAATTTTGGCAGGTTATGGACAGGGTGGCAGC 4308  
Db 4261 ATAGGCTTTGGAAAGTGTGTGGAATTTTGGCAGGTTATGGACAGGGTGGCAGC 4320  
QY 4309 GCGCTCGTGGCTTTAAGGTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAC 4368  
Db 4321 GCGCTCGTGGCTTTAAGGTCATGAGCGGAGTTCCTCCACCGAGGACCTGGTTAC 4380  
QY 4369 CTACTCCCTGCTATCCTCTCCCTGGCGCCTAGTCTCGGGTCTGTGCGCAGCGATA 4428  
Db 4381 CTACTCCCTGCTATCCTCTCCCTGGCGCCTAGTCTCGGGTCTGTGCGCAGCGATA 4440  
QY 4429 CTGCGTGGGACGCTGGGCGGAGGGGGTGTGCACTGATGAACCGGCTGATAGCG 4488  
Db 4441 CTGCGTGGGACGCTGGGCGGAGGGGGTGTGCACTGATGAACCGGCTGATAGCG 4500  
QY 4489 TTCGCTTCGCGGGTAAACACAGTCTCCCGACGCACTATGTGCTCAGAGCGAGCTGCA 4548  
Db 4501 TTCGCTTCGCGGGTAAACACAGTCTCCCGACGCACTATGTGCTCAGAGCGAGCTGCA 4560  
QY 4549 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGTCTGTGAAGAGGCTTCA 4608  
Db 4561 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGTCTGTGAAGAGGCTTCA 4620  
QY 4609 CAGTGTGATCAACAGGAGTGTCTCAGCCATGCTCGGCTCGTGGCTTAAGATGTTGG 4668  
Db 4621 CAGTGTGATCAACAGGAGTGTCTCAGCCATGCTCGGCTCGTGGCTTAAGATGTTGG 4680  
QY 4669 GATTGTGATGACAGGCTGTGATGATTTCAAGCTGGCTCAGTCCAGTCCCTGCGG 4728  
Db 4681 GATTGTGATGACAGGCTGTGATGATTTCAAGCTGGCTCAGTCCAGTCCCTGCGG 4740  
QY 4729 CGATTCCCGGGAGTCCCTTCTTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGGC 4788  
Db 4741 CGATTCCCGGGAGTCCCTTCTTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGGC 4800

QY 4789 GACGGCATCATGCAAAACACCTGCCCATGTGGAGCACAGATCACCGGACATGTAAGAAAC 4848  
DB 4801 GACGGCATCATGCAAAACACCTGCCCATGTGGAGCACAGATCACCGGACATGTAAGAAAC 4860  
QY 4849 GGTTCATGAGGATCGTGGGCTTAGGACCTGTAGTAACACGTGACATGGAAATTCGCC 4908  
DB 4861 GGTTCATGAGGATCGTGGGCTTAGGACCTGTAGTAACACGTGACATGGAAATTCGCC 4920  
QY 4909 ATTACGGGTACACACCGGCCCTCGACGGCCCTCCGGGCGCAAAATTAATCTAGGGG 4968  
DB 4921 ATTACGGGTACACACCGGCCCTCGACGGCCCTCCGGGCGCAAAATTAATCTAGGGG 4980  
QY 4969 CTGTGGCGGTGGCTGTGAGAGTACGTGAGGTTAGCGGGTGGGGATTCACACTAC 5028  
DB 4981 CTGTGGCGGTGGCTGTGAGAGTACGTGAGGTTAGCGGGTGGGGATTCACACTAC 5040  
QY 5029 GTGACGGGATGACCACTGACAACTGAAGTAAAGTCCCGTGTGAGTTCCGGCCCGAAATTC 5088  
DB 5041 GTGACGGGATGACCACTGACAACTGAAGTAAAGTCCCGTGTGAGTTCCGGCCCGAAATTC 5100  
QY 5089 TTTCAGAAAGTGGATGGGCTGGGTTGCACAGGTACGCTCCAGCGTGCANACCCCTCCTA 5148  
DB 5101 TTTCAGAAAGTGGATGGGCTGGGTTGCACAGGTACGCTCCAGCGTGCANACCCCTCCTA 5160  
QY 5149 CGGAGGAGGTACATTCCTGTTCGGGCTCAATCAATACCTGGTTGGGTTCACAGCTCCCA 5208  
DB 5161 CGGAGGAGGTACATTCCTGTTCGGGCTCAATCAATACCTGGTTGGGTTCACAGCTCCCA 5220  
QY 5209 TCGAGCCCGAACCGGACGTAGCATGTCTACTTCATGCTCACCGACCCCTCCACATTC 5268  
DB 5221 TCGAGCCCGAACCGGATGTAGCATGTCTACTTCATGCTCACCGACCCCTCCACATTC 5280  
QY 5269 ACGGCGGAGCGGCTAAGCGTAGGCTGGCCAGGGATCTCCCCCTCTTGGCCAGCTCA 5328  
DB 5281 ACGGCGGAGCGGCTAAGCGTAGGCTGGCCAGGGATCTCCCCCTCTTGGCCAGCTCA 5340  
QY 5329 TCAGTAGCCAGCTGTCTGCGCTTCCCTTGAAGCAACATGCACTACCCGTCATGACTCC 5388  
DB 5341 TCAGTAGCCAGCTGTCTGCGCTTCCCTTGAAGCAACATGCACTACCCGTCATGACTCC 5400  
QY 5389 CGGACGCTGACCTCATCGAGCCAACTCTGTGGCGGAGGAGATGGCGGGAAATTC 5448  
DB 5401 CGGACGCTGACCTCATCGAGCCAACTCTGTGGCGGAGGAGATGGCGGGAAATTC 5460  
QY 5449 ACCGCGTGGAGTCAGAAAATAAGTGTAGTAATTTTGGACTCTTTCGAGCCGCTCCAGGG 5508  
DB 5461 ACCGCGTGGAGTCAGAAAATAAGTGTAGTAATTTTGGACTCTTTCGAGCCGCTCCAGGG 5520  
QY 5509 GAGGAGATGAGAGGAAAGTATCCGTTCCGGCGGAGATCCCTCGGAGGTCCAGGAAATTC 5568  
DB 5521 GAGGAGATGAGAGGAGTATCCGTTCCGGCGGAGATCCCTCGGAGGTCCAGGAAATTC 5580  
QY 5569 CCTCGAGGATGCCATATGGCAGCGCCGGATTACACCCCTCCACTGTAGAGTCTGG 5628  
DB 5581 CCTCGAGGATGCCATATGGCAGCGCCGGATTACACCCCTCCACTGTAGAGTCTGG 5640  
QY 5629 AAGGACCCGAGTACGTCCTCCAGTGTACACGGGTGTCCATTGCGCCCTGCCAAGGCC 5688  
DB 5641 AAGGACCCGAGTACGTCCTCCAGTGTACACGGGTGTCCATTGCGCCCTGCCAAGGCC 5700  
QY 5689 CCTCGATACCACTCCACGAGGAGGACGTTGTCTGTGAGATCTACCGTGTCT 5748  
DB 5701 CCTCGATACCACTCCACGAGGAGGACGTTGTCTGTGAGATCTACCGTGTCT 5760  
QY 5749 TCTGCTTGGGGAGCTGCCAAGACCTTCGCGAGCTCCGATCGTCCGCGCTCGAC 5808  
DB 5761 TCTGCTTGGGGAGCTGCCAAGACCTTCGCGAGCTCCGATCGTCCGCGCTCGAC 5820  
QY 5809 AGCGGACGGCAACGGGCTCTCTCACAGAGCCCTCCGACGCGGAGCGCGGATCCGAC 5868  
DB 5821 AGCGGACGGCAACGGGCTCTCTCACAGAGCCCTCCGACGCGGAGCGCGGATCCGAC 5880  
QY 5869 GTTGAAGTGTACTCTCTCCATGCCCCCTTTGAGGGGAGCGCGGGGATTCGATCTCAGC 5928

DB 5881 GTTGAAGTGTACTCTCTCCATGCCCCCCCTTGGGGGAGCGCGGGGATCCCGATCTCAGC 5940  
QY 5929 GACGGTCTTGGTCTAACCGTAAGCAGAGGCTAGTAGGAGCTCGTCTGTCTGCTCGATG 5988  
DB 5941 GACGGTCTTGGTCTAACCGTAAGCAGAGGCTAGTAGGAGCTCGTCTGTCTGCTCGATG 6000  
QY 5989 TCCTACACATGGACAGGCGCCCTGATCAGGCATCGCTGCGGAGGAAACCAAGTGGCC 6048  
DB 6001 TCCTACACATGGACAGGCGCCCTGATCAGGCATCGCTGCGGAGGAAACCAAGTGGCC 6060  
QY 6049 ATCAATGACATGAGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAACTCT 6108  
DB 6061 ATCAATGACATGAGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAACTCT 6120  
QY 6109 CGCAGCGCAAGCTCGCGGAGAGAGTCACTTTTGAAGAGTCAAGTCTCGTCTGACGAC 6168  
DB 6121 CGCAGCGCAAACTCGCGGAGAGAGTCACTTTTGAAGAGTCAAGTCTCGTCTGACGAC 6180  
QY 6169 CACTACCGGAGCTGCTCAAGGAGATGAAGCGAGGCTCCACAGTTAAGCTAACTT 6228  
DB 6181 CACTACCGGAGCTGCTCAAGGAGATGAAGCGAGGCTCCACAGTTAAGCTAACTT 6240  
QY 6229 CTATCCGTGGAGGAGCTGTAAAGTGAAGCGCCCACTTGGCCAGATCTAAATTTGSC 6288  
DB 6241 CTATCCGTGGAGGAGCTGTAAAGTGAAGCGCCCACTTGGCCAGATCTAAATTTGSC 6300  
QY 6289 TATGGGCAAAAGGAGCTCGGAACTATCCAGCAAGCGGTTAAACATCGCTCCGCTG 6348  
DB 6301 TATGGGCAAAAGGAGCTCGGAACTATCCAGCAAGCGGTTAAACATCGCTCCGCTG 6360  
QY 6349 TGGAGGACTTGTCTGGAAGACACTGAGACACCAATTGACACCACTCATGGCAAAAT 6408  
DB 6361 TGGAGGACTTGTCTGGAAGACACTGAGACACCAATTGACACCACTCATGGCAAAAT 6420  
QY 6409 GAGGTTTTCTCGTCCAAACAGAGAGGCGGCGCAAGCCAGCTCGGCTTATCGTATTC 6468  
DB 6421 GAGGTTTTCTCGTCCAAACAGAGAGGCGGCGCAAGCCAGCTCGGCTTATCGTATTC 6480  
QY 6469 CCAGATTTGGGGTTCGTTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTC 6528  
DB 6481 CCAGATTTGGGGTTCGTTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTC 6540  
QY 6529 CCTCAGGCGGTGATGGGCTCTTACAGGATTCAAATCTCTCTGGACAGCGGTCGAG 6588  
DB 6541 CCTCAGGCGGTGATGGGCTCTTACAGGATTCAAATCTCTCTGGACAGCGGTCGAG 6600  
QY 6589 TTCTCTGTGAATGCTCGAAAGCGAAATGSCCTATGCGCTTTCGCAATATGACACCCGC 6648  
DB 6601 TTCTCTGTGAATGCTCGAAAGCGAAATGSCCTATGCGCTTTCGCAATATGACACCCGC 6660  
QY 6649 TGTCTTGAATCAACGGTCACTGAGATGACATCGTGTGAGGAGTCAATCTACCAATGT 6708  
DB 6661 TGTCTTGAATCAACGGTCACTGAGATGACATCGTGTGAGGAGTCAATCTACCAATGT 6720  
QY 6709 TGTGACTTGGCCCCCGGAGCCAGACAGGCCATAAGTGTGCTCACAGAGCGGCTTTACATC 6768  
DB 6721 TGTGACTTGGCCCCCGGAGCCAGACAGGCCATAAGTGTGCTCACAGAGCGGCTTTACATC 6780  
QY 6769 GGGGCCCCCTGACTAAATTTCTAAAGGCGAGAACTGCGGCTATCGCCGCTCGCGCGGAGC 6828  
DB 6781 GGGGCCCCCTGACTAAATTTCTAAAGGCGAGAACTGCGGCTATCGCCGCTCGCGCGGAGC 6840  
QY 6829 GGTCTACTGAGCACAGCTGGGTAATACCTCATGTTTACTTTGAAGGCGCTGCGGCC 6888  
DB 6841 GGTCTACTGAGCACAGCTGGGTAATACCTCATGTTTACTTTGAAGGCGCTGCGGCC 6900  
QY 6889 TGTGAGCTGGAAGCTCCAGGACTGCAAGTGTCTGATGCGGAGAGCACTTGTCTGTT 6948  
DB 6901 TGTGAGCTGGAAGCTCCAGGACTGCAAGTGTCTGATGCGGAGAGCACTTGTCTGTT 6960  
QY 6949 ATCTGTGAAGCGCGGGAGCCCAAGAGCAAGCGGAGCTTACGGGCTTTCAGGAGGCT 7008

6961 ATCTGTGAAGCGGGGACCCAAAGAGGAGCGAGCGCTACGGGCTTCACGGAGGCT 7020  
7009 ATGACTAGATACCTGCGCCGCCCTGGGGACCCGCCCAAAACAGAAATACGACTTGGAGTTG 7068  
7021 ATGACTAGATACCTGCGCCGCCCTGGGGACCCGCCCAAAACAGAAATACGACTTGGAGTTG 7080  
7069 ATAACATCATGCTCTCCCAATGTGTAGTCGGCGACGATGCTGCTGCAAAAGGTTGATC 7128  
7081 ATAACATCATGCTCTCCCAATGTGTAGTCGGCGACGATGCTGCTGCAAAAGGTTGATC 7140  
7129 TATCTCACCGGTGACCCACACACCCCTTGGCGGGCTCGTGGGAGACAGCTAGACAC 7188  
7141 TATCTCACCGGTGACCCACACCCCTTGGCGGGCTCGTGGGAGACAGCTAGACAC 7200  
7189 ACTCCAGTAACTCTCGGTGAGGAACATCATGATGTCGCCCACTTGTGGGCAAGG 7248  
7201 ACTCCAGTAACTCTCGGTGAGGAACATCATGATGTCGCCCACTTGTGGGCAAGG 7260  
7249 ATGATCCTGATGACTCATTTCTCTCATCTCTAGCTCAGGAACAACCTGGAAGAAGCC 7308  
7261 ATGATCCTGATGACTCATTTCTCTCATCTCTAGCTCAGGAACAACCTGGAAGAAGCC 7320  
7309 CTAGATTGTCAGATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTACCTCAGATC 7368  
7321 CTAGATTGTCAGATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTACCTCAGATC 7380  
7369 ATTCAAGGACTCAATGGCTTAGGCAATTTCTCTCATGATTTACTCTCAGGTGATC 7428  
7381 ATTCAAGGACTCAATGGCTTAGGCAATTTCTCTCATGATTTACTCTCAGGTGATC 7440  
7429 AATAGGTGCTCTCATGCTCAGAAACTTGGGTACCGCTTGGAGTCTCGAGACAT 7488  
7441 AATAGGTGCTCTCATGCTCAGAAACTTGGGTACCGCTTGGAGTCTCGAGACAT 7500  
7489 CGGGCCAGAAAGTCTCGCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGC 7548  
7501 CGGGCCAGAAAGTCTCGCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGC 7560  
7549 AAGTACTCTTCAACTGGGAGTAAGAACCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7608  
7561 AAGTACTCTTCAACTGGGAGTAAGAACCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7620  
7609 TCCAGTGGATTTATCCAGCTGGTCTGTTGCTGTAGCTTACAGCGGGGAGACATATATCAC 7668  
7621 TCCAGTGGATTTATCCAGCTGGTCTGTTGCTGTAGCTTACAGCGGGGAGACATATATCAC 7680  
7669 AGCTGTCTCGTGCCGACCCCGCTGTTCAATGTGTGCTACTCTCTACTTTCTGTAGGG 7728  
7681 AGCTGTCTCGTGCCGACCCCGCTGTTCAATGTGTGCTACTCTCTACTTTCTGTAGGG 7740  
7729 GTAGGCATCTATCTATCCGCCAACCGATGAACCGGGACCTTAACACTCCAGGCCAATAGG 7788  
7741 GTAGGCATCTATCTATCCGCCAACCGATGAACCGGGAGCTTAACACTCCAGGCCAATAGG 7800  
7789 CCATCTGTTTTTCTCCCTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTTTTTTTTTTTTT 7848  
7801 CCATCTGTTTTTCTCCCTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTTTTTTTTTTTTT 7860  
7849 TTTTCTCTTTTTTTTCTCTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTTTTTTTTTTTTT 7908  
7861 TTTTCTCTTTTTTTTCTCTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTTTTTTTTTTTTT 7920  
7909 CCTAGTCAGGCTAGCTGTGAAGAGTCCGTGAGCGGCTTGACTGCGAGAGTCTGATAC 7968  
7921 CCTAGTCAGGCTAGCTGTGAAGAGTCCGTGAGCGGCTTGACTGCGAGAGTCTGATAC 7980  
7969 TGGCCCTCTCTGAGATCAAGT 7989  
7981 TGGCCCTCTCTGAGATCAAGT 8001

LOCUS AX036261 8001 bp DNA linear PAT 16-NOV-2000  
DEFINITION Sequence 10 from Patent EP1043399.  
ACCESSION AX036261  
VERSION AX036261.1 GI:11225877  
KEYWORDS Hepatitis C virus  
SOURCE Hepatitis C virus  
ORGANISM Hepatitis C virus  
Viruses; sRNA positive-strand viruses, no DNA stage; Flaviviridae; Hepacivirus.  
REFERENCE 1  
AUTHORS Bartenschlager,R.D.  
TITLE Hepatitis C virus cell culture system  
JOURNAL Patent: EP 1043399-A 10 11-OCT-2000;  
BARTENSCHLAGER RALF DR (DE)  
FEATURES  
source  
1..8001  
/organism="Hepatitis C virus"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:11103"  
ORIGIN  
Query Match 99.4%; Score 7939.8; DB 6; Length 8001;  
Best Local Similarity 99.6%; Pred. No. 0;  
Matches 7972; Conservative 0; Mismatches 17; Indels 12; Gaps 1;  
QY 1 GCCAGCCCCGATTGGGGGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
Db 1 GCCAGCCCCGATTGGGGGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
QY 61 TCTTCACGACAGAAAGCGTCTAGCCATGGCGTTAGTAGTGTCTGTCGACGCTCCAGGAC 120  
Db 61 TCTTCACGACAGAAAGCGTCTAGCCATGGCGTTAGTAGTGTCTGTCGACGCTCCAGGAC 120  
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATGCCAG 180  
Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATGCCAG 180  
QY 181 GAGACCGGGTCTCTTCTTGGATCAACCGCTCAATGCTGAGATTGGCGTGGCCCC 240  
Db 181 GAGACCGGGTCTCTTCTTGGATCAACCGCTCAATGCTGAGATTGGCGTGGCCCC 240  
QY 241 GCGAGACTCTAGCCAGTAGTGTGGGTGCGAAAGGCTTGTGCTACTGCTGATAGG 300  
Db 241 GCGAGACTCTAGCCAGTAGTGTGGGTGCGAAAGGCTTGTGCTACTGCTGATAGG 300  
QY 301 GTGCTTGCAGTGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGAAATCCTAAAC 360  
Db 301 GTGCTTGCAGTGTCCCGGGAGGTCTCGTAGACCGTGCACCATGAGCAGAAATCCTAAAC 360  
QY 361 CTCAAAGAAAACCAAA-----GGGCGGCCCATGATTGAACAAGATGGATTGC 408  
Db 361 CTCAAAGAAAACCAAAACGTAACACCAAGCGCGCCCATGATTGAACAAGATGGATTGC 420  
QY 409 AGCAGGTTCTCGGCGGCTTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAG 468  
Db 421 AGCAGGTTCTCGGCGGCTTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAG 480  
QY 469 CAATCGGCTGCTGTGATGCGCGGTTCGCGCTGTGAGCGAGGGCGCCCGTCTTT 528  
Db 481 CAATCGGCTGCTGTGATGCGCGGTTCGCGCTGTGAGCGAGGGCGCCCGTCTTT 540  
QY 529 TTGTCAAGACCGACCTGTCCGGTCCCTGAATGAACCTGAGACGAGGAGCGGCTAT 588  
Db 541 TTGTCAAGACCGACCTGTCCGGTCCCTGAATGAACCTGAGGACGAGGAGCGGCTAT 600  
QY 589 CGTGGCTGGCCACGACGCGGCTTCTTGGCAGAGTGTGCTGACGCTTGTACGAGCGG 648  
Db 601 CGTGGCTGGCCACGACGCGGCTTCTTGGCAGAGTGTGCTGACGCTTGTACGAGCGG 660  
QY 649 GAAGGGACTGGCTGTATTGGGCGAAGTCCCGGGGAGGATCTCTGTCTATCTCATCTTG 708  
Db 661 GAAGGGACTGGCTGTATTGGGCGAAGTCCCGGGGAGGATCTCTGTCTATCTCATCTTG 720

QY 709 CTCCTCCGAGAAAGTATCCATCATGCTGATGCAATCGGCGGCTGCATACGCTTGATC 768  
 Db 721 CTCCTGCCGAGAAAGTATCCATCATGCTGATGCAATCGGCGGCTGCATACGCTTGATC 780  
 QY 769 CGGCTACCTTGCCCATTCGACCAACCAAGGAAACATCGCATCGAGCGAGCAGCTACTCGGA 828  
 Db 781 CGGCTACCTTGCCCATTCGACCAACCAAGGAAACATCGCATCGAGCGAGCAGCTACTCGGA 840  
 QY 829 TGGAAAGCGGCTCTTCGATCAGGATGATCTGGAAGAGAGATCAGGGGCTCGGCCAG 888  
 Db 841 TGGAAAGCGGCTCTTCGATCAGGATGATCTGGAAGAGAGATCAGGGGCTCGGCCAG 900  
 QY 889 CGGACTCTTCGCAAGGCTCAAGGCGCATCGCCGAGCGGAGGATCTCGCTGACCC 948  
 Db 901 CGGAACTGTTCGCCAGGCTCAAGGCGCATCGCCGAGCGGAGATCTCGCTGAGCCC 960  
 QY 949 ATGGCGATGCTGCTTCGCGAATATCATGCTGGAATAATGGCGCTTTCTGGATTCATCG 1008  
 Db 961 ATGGCGATGCTGCTTCGCGAATATCATGCTGGAATAATGGCGCTTTCTGGATTCATCG 1020  
 QY 1009 ACTGTGGCGGCTGGGTGTGGCGGACCGGTATCAGGACATAGCGTTGGCTACCGTGATA 1068  
 Db 1021 ACTGTGGCGGCTGGGTGTGGCGGACCGGTATCAGGACATAGCGTTGGCTACCGTGATA 1080  
 QY 1069 TTGCTGAAGAGCTTCGCGCGAATCGGCTGACCGCTTCCTGCTGTTTACGGTATCGCG 1128  
 Db 1081 TTGCTGAAGAGCTTCGCGCGAATCGGCTGACCGCTTCCTGCTGTTTACGGTATCGCG 1140  
 QY 1129 CTCGCGATTCGAGCGCATCGCTTCTATCGCTTCTTTCGAGTTCCTTCGAGTTTAAA 1188  
 Db 1141 CTCGCGATTCGAGCGCATCGCTTCTATCGCTTCTTTCGAGTTCCTTCGAGTTTAAA 1200  
 QY 1189 CAGACCAACAGGTTTCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCGCCCCCT 1248  
 Db 1201 CAGACCAACAGGTTTCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCGCCCCCT 1260  
 QY 1249 AACGTTACTGGCGGAAGCGCTTGAATAAGCGCGGTGCGGTTTGTCTATATGTTATT 1308  
 Db 1261 AACGTTACTGGCGGAAGCGCTTGAATAAGCGCGGTGCGGTTTGTCTATATGTTATT 1320  
 QY 1309 TCCACCATATTCGCTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGCTCTTTG 1368  
 Db 1321 TCCACCATATTCGCTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGCTCTTTG 1380  
 QY 1369 ACGACATTCCTAGGGGCTTTTCCCTCTCGCAGAAAGGAAATCGAAGGCTGTGTAATGTC 1428  
 Db 1381 ACGACATTCCTAGGGGCTTTTCCCTCTCGCAGAAAGGAAATCGAAGGCTGTGTAATGTC 1440  
 QY 1429 GTGAAGGAAGCAGTTTCTCTGGAAGCTTCTTGAAGACAAACAACTGCTAGCGACCTT 1488  
 Db 1441 GTGAAGGAAGCAGTTTCTCTGGAAGCTTCTTGAAGACAAACAACTGCTAGCGACCTT 1500  
 QY 1489 TCGAGCGAGCGAAACCCCCCACTCGGCAAGGTCCTCTCGGCCCAAAAGCCACGTGTA 1548  
 Db 1501 TCGAGCGAGCGAAACCCCCCACTCGGCAAGGTCCTCTCGGCCCAAAAGCCACGTGTA 1560  
 QY 1549 TAAGATACCTGCAAGGCGGCAACCCCACTCGGCAAGGTCCTCTCGGCCCAAAAGCCACGTG 1608  
 Db 1561 TAAGATACCTGCAAGGCGGCAACCCCACTCGGCAAGGTCCTCTCGGCCCAAAAGCCACGTG 1620  
 QY 1609 GAAAGGTCAATGGCTCTCTTCAAGCGTATTCAAAGGGGCTGAAGGATCCCGAGAG 1668  
 Db 1621 GAAAGGTCAATGGCTCTCTTCAAGCGTATTCAAAGGGGCTGAAGGATCCCGAGAG 1680  
 QY 1669 GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTACATGCTTTAG 1728  
 Db 1681 GTACCCCATTTGATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTACATGCTTTAG 1740  
 QY 1729 TCGAGGTTAAAAAAGCTGTAGGCCCCCGAAGCAGCGGAGCTGTTTTCTTTTCAAAAA 1788  
 Db 1741 TCGAGGTTAAAAAAGCTGTAGGCCCCCGAAGCAGCGGAGCTGTTTTCTTTTCAAAAA 1800  
 QY 1789 CACGATAATACCATGGCGCTTATTAGGCGCTACTCCCAACAGACGCGAGGCTACTTGGC 1848

Db 1801 CACGATAATACCATGGCGCTTATTAGGCGCTACTCCCAACAGACGCGAGGCTACTTGGC 1860  
 QY 1849 TGCATCATCTAGCTCAGAGCGGACAGGAAACAGGTCGAGGGGAGGTCCAAAGTG 1908  
 Db 1861 TGCATCATCTAGCTCAGAGCGGACAGGAAACAGGTCGAGGGGAGGTCCAAAGTG 1920  
 QY 1909 GTCTCCACCGCAACAAATCTTTCTTGGGACCTGCGTCAATGGGCTGTGTGGACTGTC 1968  
 Db 1921 GTCTCCACCGCAACAAATCTTTCTTGGGACCTGCGTCAATGGGCTGTGTGGACTGTC 1980  
 QY 1969 TATCATGTGCGCGCTCAAGAACCTTCCGCGCCCAAGGGGCCAATCACCCAAATGTAC 2028  
 Db 1981 TATCATGTGCGCGCTCAAGAACCTTCCGCGCCCAAGGGGCCAATCACCCAAATGTAC 2040  
 QY 2029 ACCAATGTGACAGGACCTCTGTCGGCTGGCAAGCGCCCGCGGCGCTTCTTGTACA 2088  
 Db 2041 ACCAATGTGACAGGACCTCTGTCGGCTGGCAAGCGCCCGCGGCGCTTCTTGTACA 2100  
 QY 2089 CCATGCACTCGGCGAGCTCGGACCTTTTACCTTGGTCACGAGCATGCGGATGTCATTCG 2148  
 Db 2101 CCATGCACTCGGCGAGCTCGGACCTTTTACCTTGGTCACGAGCATGCGGATGTCATTCG 2160  
 QY 2149 GTGCGCCGCGGCGGACAGCAGGCGGAGCCTACTCTCCCCAGGCGCGCTCTCTACTTGG 2208  
 Db 2161 GTGCGCGGCGGCGGACAGCAGGCGGAGCCTACTCTCCCCAGGCGCGCTCTCTACTTGG 2220  
 QY 2209 AAGGCTCTTTCGGGCGGTCCTGCTGTCGCCCTCGGGGACGCTGTGGGCACTTTTCGG 2268  
 Db 2221 AAGGCTCTTTCGGGCGGTCCTGCTGTCGCCCTCGGGGACGCTGTGGGCACTTTTCGG 2280  
 QY 2269 GCTGCGGCTGTCACCCGAGGGGTTCCGAAGGGGCTGACTTTGTATACCGTTCGAGTCTATG 2328  
 Db 2281 GCTGCGGCTGTCACCCGAGGGGTTCCGAAGGGGCTGACTTTGTATACCGTTCGAGTCTATG 2340  
 QY 2329 GAAACCACTATGCGGCTCCCGGCTCTTACGGACAACTCGTCCCTTCGCGCGCTACCGGAG 2388  
 Db 2341 GGAACCACTATGCGGCTCCCGGCTCTTACGGACAACTCGTCCCTTCGCGCGCTACCGGAG 2400  
 QY 2389 ACATTCAGGTGGGCCATCTACAGCCCTACTGTAGCGGCAAGACACTAAGTGTGCG 2448  
 Db 2401 ACATTCAGGTGGGCCATCTACAGCCCTACTGTAGCGGCAAGACACTAAGTGTGCG 2460  
 QY 2449 GCTGGTATGAGCCCAAGGCTATAGGCTGTCTGCTGAACCGCTCGTTCGCGCGCAC 2508  
 Db 2461 GCTGGTATGAGCCCAAGGCTATAGGCTGTCTGCTGAACCGCTCGTTCGCGCGCAC 2520  
 QY 2509 CTAGGTTTCGGGCGGTATATGTCTAAGGCAATGCTATCGACCTTAAACATCAGAAACCGGG 2568  
 Db 2521 CTAGGTTTCGGGCGGTATATGTCTAAGGCAATGCTATCGACCTTAAACATCAGAAACCGGG 2580  
 QY 2569 GTAAGGACCATCAACACCGGTCGCCCATACGTAATCTGACTATGGAAGGAGGCTTCTTGGC 2628  
 Db 2581 GTAAGGACCATCAACACCGGTCGCCCATACGTAATCTGACTATGGAAGGAGGCTTCTTGGC 2640  
 QY 2629 GACGCTGCTGCTCGGGGCGCTTACACATCAATATGATGATGAGTGCACACTCAACT 2688  
 Db 2641 GACGCTGCTGCTCGGGGCGCTTACACATCAATATGATGATGAGTGCACACTCAACT 2700  
 QY 2689 GACTCGACCATATCTCTGGGCGATCGGACAGTCTCTGACCAAGCGGAGGAGCTGGAGCG 2748  
 Db 2701 GACTCGACCATATCTCTGGGCGATCGGACAGTCTCTGACCAAGCGGAGGAGCTGGAGCG 2760  
 QY 2749 GACTGCTGCTGCTCGGACACCGCTACGCTCGGGATCGGTCAACGTCGCACTCAAAAC 2808  
 Db 2761 GACTGCTGCTGCTCGGACACCGCTACGCTCGGGATCGGTCAACGTCGCACTCAAAAC 2820  
 QY 2809 ATCGAGGAGGTGGCTCTCTCCAGCACTCGGAAATCCCTTTTATGCGAAAGCCATCCCC 2868  
 Db 2821 ATCGAGGAGGTGGCTCTCTCCAGCACTCGGAAATCCCTTTTATGCGAAAGCCATCCCC 2880  
 QY 2869 ATCGAGACCATCAAGGGGCGGAGCCTCTATTTCTGCCATTCCTCAAGAAATGTGAT 2928



Db 2881 ATCGAGCCATCAAGGGGGAGCGACCTCATTTTCTGCCATTTCCAGAGAAATGTGAT 2940  
QY 2929 GAGCTCGCGCGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCAATATACCGGGCCCTT 2988  
Db 2941 GAGCTCGCGCGAAGCTGTCCGGCTCGGACTCAATGCTGTAGCAATATACCGGGCCCTT 3000  
QY 2989 GATGATCCGTATACCAATAGCGAGAGCGTCAATGCTGTAGCAAGCGGACTCTAAATG 3048  
Db 3001 GATGATCCGTATACCAATAGCGAGAGCGTCAATGCTGTAGCAAGCGGACTCTAAATG 3060  
QY 3049 ACGGGCTTTACCGCGGATTTCCGACTCAGTCATCGAATGCAATGTCATGCCAGACA 3108  
Db 3061 ACGGGCTTTACCGCGGACTTCGACTCAGTCATCGAATGTCATGCCAGACA 3120  
QY 3109 GTCGACTTCAGCTGAGCCGCACTTCAANTGAGACGACGCGTGCACAAAGACGG 3168  
Db 3121 GTCGACTTCAGCTGAGCCGCACTTCAANTGAGACGACGCGTGCACAAAGACGG 3180  
QY 3169 GTGTACGCTCGCAGCGCGAGGACGACTGGTAGGGGAGGATGGGCATTTACAGTTT 3228  
Db 3181 GTGTACGCTCGCAGCGCGAGGACGACTGGTAGGGGAGGATGGGCATTTACAGTTT 3240  
QY 3229 GTGACTTCCAGGAAACGGCCCTCGGCATGTTCCGATTCCTCGGTTCTGTGCGAGTCTAT 3288  
Db 3241 GTGACTTCCAGGAAACGGCCCTCGGCATGTTCCGATTCCTCGGTTCTGTGCGAGTCTAT 3300  
QY 3289 GAGCGGGCTGTCTTGTGTAAGAGTCAACCGCCGCGAGACCTCACTAGTTAGTTGCGGGT 3348  
Db 3301 GAGCGGGCTGTCTTGTGTAAGAGTCAACCGCCGCGAGACCTCACTAGTTAGTTGCGGGT 3360  
QY 3349 TACCTAAACACACCGGGTTGCCGTCGCCAGGACCATCTGGAGTCTGGGAGAGCGTC 3408  
Db 3361 TACCTAAACACACCGGGTTGCCGTCGCCAGGACCATCTGGAGTCTGGGAGAGCGTC 3420  
QY 3409 TTTACAGGCTCACCCACATAGACGCCCATTTCTGTCCAGACTAAGCAGGACGAGAC 3468  
Db 3421 TTTACAGGCTCACCCACATAGACGCCCATTTCTGTCCAGACTAAGCAGGACGAGAC 3480  
QY 3469 AACTTCCCTTACTGTGTAGATACAGGCTACGGTGTGGCCAGGCTCAGGCTCCACCT 3528  
Db 3481 AACTTCCCTTACTGTGTAGATACAGGCTACGGTGTGGCCAGGCTCAGGCTCCACCT 3540  
QY 3529 CCATCGTGGGACCAATGTGGAGTCTCTATACGGCTAAGCTGACAGTGCACGGGCA 3588  
Db 3541 CCATCGTGGGACCAATGTGGAGTGTCTATACGGCTAAGCTGACAGTGCACGGGCA 3600  
QY 3589 ACGCCCTGCTGTATAGGCTGGAGCGGTTCAAAACGAGTTACTACACACACCCCAT 3648  
Db 3601 ACGCCCTGCTGTATAGGCTGGAGCGGTTCAAAACGAGTTACTACACACACCCCAT 3660  
QY 3649 ACCAAATACATATGGCATGTGTCGGCTGACCTGGAGTGTCTACGAGACCTGGTG 3708  
Db 3661 ACCAAATACATATGGCATGTGTCGGCTGACCTGGAGTGTCTACGAGACCTGGTG 3720  
QY 3709 CTGTTAGGCGGAGTCTTACGAGTCTGGCGGTATTTGCTGTACACAGCAGCGTGT 3768  
Db 3721 CTGTTAGGCGGAGTCTTACGAGTCTGGCGGTATTTGCTGTACACAGCAGCGTGT 3780  
QY 3769 ATTGTGGGAGGATCATTTGTCGGAAAGCCGGCCATCATTTCCGACAGGGAAGTCTTT 3828  
Db 3781 ATTGTGGGAGGATCATTTGTCGGAAAGCCGGCCATCATTTCCGACAGGGAAGTCTTT 3840  
QY 3829 TACCGGAGTTTGATGAGATGGAAGATGCGCTCACACTCCCTTACATCGAACAGGGA 3888  
Db 3841 TACCGGAGTTTGATGAGATGGAAGATGCGCTCACACTCCCTTACATCGAACAGGGA 3900  
QY 3889 ATGACGTCGCCAAACAAATTAACAGAGGCAATCGGTTGTGTGCAAAACAGCCACAAG 3948  
Db 3901 ATGACGTCGCCAAACAAATTAACAGAGGCAATCGGTTGTGTGCAAAACAGCCACAAG 3960  
QY 3949 CAAGCGGAGGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCTCGAAGCTTCTGG 4008  
Db 3961 CAAGCGGAGGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCTCGAAGCTTCTGG 4020

QY 4009 GCGAAGCATATGTGGAATTTTCATACGGGATACAATATTTAGCAGGCTTGTCCACTCTG 4068  
Db 4021 GCGAAGCATATGTGGAATTTTCATACGGGATACAATATTTAGCAGGCTTGTCCACTCTG 4080  
QY 4069 CTGTGCAACCCCGGATAGCATCACTGATGGGATTCACAGCTCTATCACCAGCGGCTC 4128  
Db 4081 CTGTGCAACCCCGGATAGCATCACTGATGGGATTCACAGCTCTATCACCAGCGGCTC 4140  
QY 4129 ACCACCAACATACCTCTCTTTTAAATCCTTGGGGGATGGGTGGCCGCCCACTTGTCT 4188  
Db 4141 ACCACCAACATACCTCTCTTTTAAATCCTTGGGGGATGGGTGGCCGCCCACTTGTCT 4200  
QY 4189 CTCTCCAGCGCTGTCTCTCTCTTGTGAGGCGCGGATCGCTGGAGCGGCTGTGGCAGC 4248  
Db 4201 CTCTCCAGCGCTGTCTCTCTCTTGTGAGGCGCGGATCGCTGGAGCGGCTGTGGCAGC 4260  
QY 4249 ATAGGCTTGGGAAGTGTCTTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGC 4308  
Db 4261 ATAGGCTTGGGAAGTGTCTTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGC 4320  
QY 4309 GGGCTCGTGGCTTTAAGGTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAAC 4368  
Db 4321 GCGCTGTGGCTTTAAGGTCATGAGCGCGAGTTCGCCCTCCACCGAGGACCTGGTTAAC 4380  
QY 4369 CTACTCCCTGTCTATCT 4428  
Db 4381 CTACTCCCTGTCTATCT 4440  
QY 4429 CTGCGTGGCAGCTGGGCGCCAGGGAGGGGCTGTGTCAGTGGATGAACCGGCTGATAGC 4488  
Db 4441 CTGCGTGGCAGCTGGGCGCCAGGGAGGGGCTGTGTCAGTGGATGAACCGGCTGATAGC 4500  
QY 4489 TTGCTTTCGGGGTAAACACAGCTCTCCCCACGCACTATGTGCTCAGAGCGGCTGCA 4548  
Db 4501 TTGCTTTCGGGGTAAACACAGCTCTCCCCACGCACTATGTGCTCAGAGCGGCTGCA 4560  
QY 4549 GCACTGTCTCAGATCT 4608  
Db 4561 GCACTGTCTCAGATCT 4620  
QY 4609 CAGTGGATCAACGAGGACTGCTCCAGCCATGCTCGGCTCGTGGCTAGAGATGTTGG 4668  
Db 4621 CAGTGGATCAACGAGGACTGCTCCAGCCATGCTCGGCTCGTGGCTAGAGATGTTGG 4680  
QY 4669 GATTGGATATGACGCGTGTGATTTCAAGACCTCGCTCCAGTCCCAAGCTCTCTGCCG 4728  
Db 4681 GATTGGATATGACGCGTGTGATTTCAAGACCTCGCTCCAGTCCCAAGCTCTCTGCCG 4740  
QY 4729 CGATTGCGGGAGTCCCT 4788  
Db 4741 CGATTGCGGGAGTCCCT 4800  
QY 4789 GACGCGATCATCAACACACCTGCGCATCTGGAGACAGATCACCGGACATGTGAAAAAC 4848  
Db 4801 GACGCGATCATCAACACACCTGCGCATCTGGAGACAGATCACCGGACATGTGAAAAAC 4860  
QY 4849 GGTTCATGAGGATCGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAAATTTCCCT 4908  
Db 4861 GGTTCATGAGGATCGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAAATTTCCCT 4920  
QY 4909 ATTACCGGTAACACCGGGCCCTTGACGCGCTCTCCCGGCCCAATTTATTCTAGGGCG 4968  
Db 4921 ATTACCGGTAACACCGGGCCCTTGACGCGCTCTCCCGGCCCAATTTATTCTAGGGCG 4980  
QY 4969 CTGTGGCGGTGCTCTGAGGAGTACGTGGAGTGTACCGGGTGGGGATTTCCACTAC 5028  
Db 4981 CTGTGGCGGTGCTCTGAGGAGTACGTGGAGTGTACCGGGTGGGGATTTCCACTAC 5040  
QY 5029 GTGAGGGGATGACCACTGACAACTGAAAGTGGCGGTGTGAGTCCGGCCCCCGGAATTC 5088  
Db 5041 GTGAGGGGATGACCACTGACAACTGAAAGTGGCGGTGTGAGTCCGGCCCCCGGAATTC 5100

QY 5089 TTCAAGAGTGGATGGGGTGGGTTGACAGGTACGCTCCAGCGTGCAGAACCCCTCCTA 5148  
Db 5101 TTCAAGAGTGGATGGGGTGGGTTGACAGGTACGCTCCAGCGTGCAGAACCCCTCCTA 5160  
QY 5149 CGGAGGAGGTACATTCCTGGTGGGGTCAATCAATACCTGGTGGGTGCACAGTCCCA 5208  
Db 5161 CGGAGGAGGTACATTCCTGGTGGGGTCAATCAATACCTGGTGGGTGCACAGTCCCA 5220  
QY 5209 TGGAGCCGGAACCGGAGTACAGTACGCTCACTTCATGCTCACCGAGCCCTCCACATTT 5269  
Db 5221 TGGAGCCGGAACCGGAGTACAGTACGCTCACTTCATGCTCACCGAGCCCTCCACATTT 5280  
QY 5269 ACGGCGGAGACGGGTAAAGCTAGGCTGGCCAGGGGATCTCCGCCCTCTTGGCCAGCTCA 5328  
Db 5281 ACGGCGGAGACGGGTAAAGCTAGGCTGGCCAGGGGATCTCCGCCCTCTTGGCCAGCTCA 5340  
QY 5329 TCAGTACGCCAGCTGTCTGGCCCTTCTTGAAGGCAACATGCACTACCCGCTCATGACTCC 5388  
Db 5341 TCAGTACGCCAGCTGTCTGGCCCTTCTTGAAGGCAACATGCACTACCCGCTCATGACTCC 5400  
QY 5389 CGGAGCGGTGACCTCATCGAGGCCAACCTCTGTGGCGGAGGATGGGGGGAGACATC 5448  
Db 5401 CGGAGCGGTGACCTCATCGAGGCCAACCTCTGTGGCGGAGGATGGGGGGAGACATC 5460  
QY 5449 ACCCGCGTGGAGTCAGAAATAAGTAGTAAATTTTGGACTCTTTTCGAGCGGCTCCAAGCG 5508  
Db 5461 ACCCGCGTGGAGTCAGAAATAAGTAGTAAATTTTGGACTCTTTTCGAGCGGCTCCAAGCG 5520  
QY 5509 GAGGAGGATGAGAGGGAATGTCGTTCCGGCGGAGATCTTTCGCGAGGTCCAGGAAATTC 5568  
Db 5521 GAGGAGGATGAGAGGGAATGTCGTTCCGGCGGAGATCTTTCGCGAGGTCCAGGAAATTC 5580  
QY 5569 CCTGAGCGATGCCATATGGCGACGCGCGGATACAAACCTCCACTTTAGAGTCTGG 5628  
Db 5581 CCTGAGCGATGCCATATGGCGACGCGCGGATACAAACCTCCACTTTAGAGTCTGG 5640  
QY 5629 AAGGACCCGGACTACGTCCTCCAGTGGTACACGGGTGTCAATTCGCCGCTGCCAAGGCC 5688  
Db 5641 AAGGACCCGGACTACGTCCTCCAGTGGTACACGGGTGTCAATTCGCCGCTGCCAAGGCC 5700  
QY 5689 CCTCGATACCACTTCCAAGGAGGAGGACGTTGTCCTGTGAGATCTACCGTGTCT 5748  
Db 5701 CCTCGATACCACTTCCAAGGAGGAGGACGTTGTCCTGTGAGATCTACCGTGTCT 5760  
QY 5749 TCTGCTTGGCGGAGCTCGCCACAAAGACCTTCGACGCTCGATTCGTCGCGCTCGAC 5808  
Db 5761 TCTGCTTGGCGGAGCTCGCCACAAAGACCTTCGACGCTCGATTCGTCGCGCTCGAC 5820  
QY 5809 AGCGGACGGCAACGGGCTCTCCTGACAGGCTCCGACGAGCGGCGGATCCGAC 5868  
Db 5821 AGCGGACGGCAACGGGCTCTCCTGACAGGCTCCGACGAGCGGCGGATCCGAC 5880  
QY 5869 GTTGAATGCTACTCTCTCATGCCCTTTCGAGGGGAGCGGGGATCCCGATCTCAGC 5928  
Db 5881 GTTGAATGCTACTCTCTCATGCCCTTTCGAGGGGAGCGGGGATCCCGATCTCAGC 5940  
QY 5929 GACGGCTCTGCTTACCGTAAAGGAGGAGCTAGTAGGAGCTGCTGCTGCTGCTGATG 5988  
Db 5941 GACGGCTCTGCTTACCGTAAAGGAGGAGCTAGTAGGAGCTGCTGCTGCTGCTGATG 6000  
QY 5989 TCCTACATGAGCAGGCGCTCTGATCAGGCTATGCTGGGGAGGAAACCAAGTGGCC 6048  
Db 6001 TCCTACATGAGCAGGCGCTCTGATCAGGCTATGCTGGGGAGGAAACCAAGTGGCC 6060  
QY 6049 ATCAATGACTGAGCACTCTTTCGCTCGCTCACCAAACTTGGTCTATGCTACAACTCT 6108  
Db 6061 ATCAATGACTGAGCACTCTTTCGCTCGCTCACCAAACTTGGTCTATGCTACAACTCT 6120  
QY 6109 CGAGCGCAAGCTGCGCAGAAAGGTTCACCTTTGACAGACTGACAGTCTCTGACGAC 6168  
Db 6121 CGAGCGCAAGCTGCGCAGAAAGGTTCACCTTTGACAGACTGACAGTCTCTGACGAC 6180  
QY 6169 CACTACCGGAGCTGCTCAAGGAGATGAAGCGGCTCCACAGTTAAGGCTAAACTT 6228

Db 6181 CACTACCGGAGCTGCTCAAGGAGATGAAGCGGAGGCTCCACAGTTAAGCTAAACTT 6240  
QY 6229 CTATCCGTGGAGGAAGCCTGTAAGCTGACGCCCCACATTGGCGAGATCTAAATTTGGC 6288  
Db 6241 CTATCCGTGGAGGAAGCCTGTAAGCTGACGCCCCACATTGGCGAGATCTAAATTTGGC 6300  
QY 6289 TATGGGCAAAAGGAGCTCGGAACTTATCCAGCAGGCGGTTTACCACTCGCTCCGCTG 6348  
Db 6301 TATGGGCAAAAGGAGCTCGGAACTTATCCAGCAGGCGGTTTAAACCATCGCTCCGCTG 6360  
QY 6349 TGGAGGACTTGTCTGGAAGACACTGAGACACCAATTGACACCATCATGGCAAAAT 6408  
Db 6361 TGGAGGACTTGTCTGGAAGACACTGAGACACCAATTGACACCATCATGGCAAAAT 6420  
QY 6409 GAGTCTTTCGCTCCAAACAGAGAGAGGGGGCGCAAGCGAGCTCGCTTATCTGATTC 6468  
Db 6421 GAGTCTTTCGCTCCAAACAGAGAGAGGGGGCGCAAGCGAGCTCGCTTATCTGATTC 6480  
QY 6469 CCAGATTTGGGGTTCGTTGTCGAGAAATGGCCCTTTAGATGTGTCTCCACCTC 6528  
Db 6481 CCAGATTTGGGGTTCGTTGTCGAGAAATGGCCCTTTAGATGTGTCTCCACCTC 6540  
QY 6529 CCTCAGGCGGTGATGGGCTCTTTCATACGGAATTCCTGGAAGCGGCTCGAG 6588  
Db 6541 CCTCAGGCGGTGATGGGCTCTTTCATACGGAATTCCTGGAAGCGGCTCGAG 6600  
QY 6589 TTCTGTTGAAATGCTTGGAAAGGAAATGGCCCTTATGGGCTTCGATATGACACCGC 6648  
Db 6601 TTCTGTTGAAATGCTTGGAAAGGAAATGGCCCTTATGGGCTTCGATATGACACCGC 6660  
QY 6649 TGTGTTGACTCAACGGTCACTGAGATGACATCGTGTGAGGAGTCAATCTACCAATGT 6708  
Db 6661 TGTGTTGACTCAACGGTCACTGAGATGACATCGTGTGAGGAGTCAATCTACCAATGT 6720  
QY 6709 TGTGACTTTGGCCCCCGAAGCCAGACAGGCCATAGGTCGCTCACAGAGCGGCTTTACATC 6768  
Db 6721 TGTGACTTTGGCCCCCGAAGCCAGACAGGCCATAGGTCGCTCACAGAGCGGCTTTACATC 6780  
QY 6769 GGGGGCCCCCTGACTAATTTCTAAAGGCGAGAACTGGCGGCTATCGCGGTCGCGCGAGC 6828  
Db 6781 GGGGGCCCCCTGACTAATTTCTAAAGGCGAGAACTGGCGGCTATCGCGGTCGCGCGAGC 6840  
QY 6829 GGTGACTGACGACCGAGCTGGGTAATACCTCATGTTACTTTGAAGCGCGCTCGCGCC 6888  
Db 6841 GGTGACTGACGACCGAGCTGGGTAATACCTCATGTTACTTTGAAGCGCGCTCGCGCC 6900  
QY 6889 TGTGAGCTGCGAAGCTCCAGGACTGCAAGTGTCTGATGCGGAGACGACCTTTCGTT 6948  
Db 6901 TGTGAGCTGCGAAGCTCCAGGACTGCAAGTGTCTGATGCGGAGACGACCTTTCGTT 6960  
QY 6949 ATCTGTGAAAGCGGGGACCCAGAGGACGAGGGAGCCTACGGGCTTACGGAGGT 7008  
Db 6961 ATCTGTGAAAGCGGGGACCCAGAGGACGAGGGAGCCTACGGGCTTACGGAGGT 7020  
QY 7009 ATGACTAGATCTCTGCCCCCTGGGACCCGCCAAACAGAAATAGACTTGGAGTTG 7068  
Db 7021 ATGACTAGATCTCTGCCCCCTGGGACCCGCCAAACAGAAATAGACTTGGAGTTG 7080  
QY 7069 ATAACTCATGCTCTCTCAATGTGTCACTGCTGCGCAAGTGTGATCTGGCAAAAGGTGTAC 7128  
Db 7081 ATAACTCATGCTCTCTCAATGTGTCACTGCTGCGCAAGTGTGATCTGGCAAAAGGTGTAC 7140  
QY 7129 TATCTCACCGGTGACCCACCCCTTGGCGGGTGTGGGAGACGACTAGACAC 7188  
Db 7141 TATCTCACCGGTGACCCACCCCTTGGCGGGTGTGGGAGACGACTAGACAC 7200  
QY 7189 ACTCCAGTCAATCTCTGGCTAGGCAACATCATCATATGATGCGCCCACTTTGGGCAAG 7248  
Db 7201 ACTCCAGTCAATCTCTGGCTAGGCAACATCATCATATGATGCGCCCACTTTGGGCAAG 7260  
QY 7249 ATGATCTGATGACTCTTCTCTCATCTCTCTAGCTCAGGAACTTGAAGGCTAAAGCC 7308

Db 7261 ATGATCTGATGACTCATTTCTTCTCCATCTCTTAGCTCAGGAACAACCTTGAAAGACC 7320

QY 7309 CTAGATTGTGATGATCTAGGGGCTGTACTCCATTGAGCACTTGACCTACCTCAGATC 7368

Db 7321 CTAGATTGTGATGATCTAGGGGCTGTACTCCATTGAGCACTTGACCTACCTCAGATC 7380

QY 7369 ATTCAACGACTTCCATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7428

Db 7381 ATTCAACGACTTCCATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7440

QY 7429 AATAGGTGGCTTTCATGCTCAGAAACTTGGGTACCGCCCTTGGCAGTCTCGAGACAT 7488

Db 7441 AATAGGTGGCTTTCATGCTCAGAAACTTGGGTACCGCCCTTGGCAGTCTCGAGACAT 7500

QY 7489 CGGGCCAGAGTGTCCGCGTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGC 7548

Db 7501 CGGGCCAGAGTGTCCGCGTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGC 7560

QY 7549 AAGTACTCTTCAACTGGGCAAGTAAGCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7608

Db 7561 AAGTACTCTTCAACTGGGCAAGTAAGCAAGCTCAAACTCACTCCAAATCCCGGCTGG 7620

QY 7609 TCCAGTGTGATTTATCAGCTGGTTGGTTGTCTGGTTACAGCGGGGAGACATATATCAC 7668

Db 7621 TCCAGTGTGATTTATCAGCTGGTTGGTTGTCTGGTTACAGCGGGGAGACATATATCAC 7680

QY 7669 AGCTGTCTGTCGCGGACCGCTGGTTGATGTGGTGGCTACTCTCTACTTTCTGTAGGG 7728

Db 7681 AGCTGTCTGTCGCGGACCGCTGGTTGATGTGGTGGCTACTCTCTACTTTCTGTAGGG 7740

QY 7729 GTAGGCACTATCTATCTCCCAACCGGATGAACGGGACCTTAAACACTCCAGGCCAATAGG 7788

Db 7741 GTAGGCACTATCTATCTCCCAACCGGATGAACGGGAGCTTAAACACTCCAGGCCAATAGG 7800

QY 7789 CCATCTGTTTTTTTCCCTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7848

Db 7801 CCATCTGTTTTTTTCCCTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860

QY 7849 TTTTCTCCTTTTTTTTCCCTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7908

Db 7861 TTTTCTCCTTTTTTTTCCCTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7920

QY 7909 CCTAGTCAGGCTGAGTGTGAAGGCTCCGTCGAGCCGCTTCACTGCAGAGAGTCTCATAC 7968

Db 7921 CCTAGTCAGGCTGAGTGTGAAGGCTCCGTCGAGCCGCTTCACTGCAGAGAGTCTCATAC 7980

QY 7969 TGGCCTCTCTGCAATCAAGT 7989

Db 7981 TGGCCTCTCTGCAATCAAGT 8001

## RESULT 14

AX937622  
LOCUS 12305 bp DNA linear PAT 06-JAN-2004  
DEFINITION Sequence 2 from Patent WO03091439.  
ACCESSION AX937622  
VERSION AX937622.1 GI:40713673  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
unclassified.

## REFERENCE

1  
Duggal, R. A., Patrick, A. K., Zhang, J. A. and Zhao, W. A.  
Reporter-selectable hepatitis c virus replicon  
Patent: WO 03091439-A 2 06-NOV-2003;  
PFIZER INC. (US)

## FEATURES

source  
1..12305  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
/note="A cell line wherein the nucleic acid molecule is a  
self replicating RNA molecule."

## ORIGIN

Query Match 95.3%; Score 7610.4; DB 6; Length 12305;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7611; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 378 GGGCGGCCATGATTGAACAAGATGATTCCAGCAGGTTCTCCGGCCGCTTGGGTGGAG 437

Db 1370 GGGCGGCCATGATTGAACAAGATGATTCCAGCAGGTTCTCCGGCCGCTTGGGTGGAG 1429

QY 438 AGCTATTTCGGCTATGACTGGGCACACACAAATCGGCTGCTCTGATGCGCGCTGTC 497

Db 1430 AGCTATTTCGGCTATGACTGGGCACACAAATCGGCTGCTCTGATGCGCGCTGTC 1489

QY 498 CGGCTGTCAAGCAGGGGGCGCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTG 557

Db 1490 CGGCTGTCAAGCAGGGGGCGCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTG 1549

QY 558 AATGAATCGACAGCAGGCGCGGCTATCGTGGCTGGCCACGACGGCGCTTCCCTGC 617

Db 1550 AATGAATCGACAGCAGGCGCGGCTATCGTGGCTGGCCACGACGGCGCTTCCCTGC 1609

QY 618 GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGGAGTG 677

Db 1610 GCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGGAGTG 1669

QY 678 CCGGGCAGGATCTCCTGTCACTCACTTGTCTCTCCGAGAAAGTATCCATCATGGCT 737

Db 1670 CCGGGCAGGATCTCCTGTCACTCACTTGTCTCTCCGAGAAAGTATCCATCATGGCT 1729

QY 738 GATGCAATGGCGGCTGATAGCTGTGATCCGGCTACCTGCCCATTCGACACCAAGCG 797

Db 1730 GATGCAATGGCGGCTGATAGCTGTGATCCGGCTACCTGCCCATTCGACACCAAGCG 1789

QY 798 AAACATCGCATCGAGCAGCACTACTCGATGGAAGCGGCTCTGTCGATCAGGATGAT 857

Db 1790 AAACATCGCATCGAGCAGCACTACTCGATGGAAGCGGCTCTGTCGATCAGGATGAT 1849

QY 858 CTGGACGAAGAGCATCAGGGGCTCGGCCAGCGAACTGTTCGCCAGGCTCAAGGCGCG 917

Db 1850 CTGGACGAAGAGCATCAGGGGCTCGGCCAGCGAACTGTTCGCCAGGCTCAAGGCGCG 1909

QY 918 ATGCCCGCAGCGGAGATCTCGTGTGACCATGGCGATGCTGCTTGGCGAATATCATG 977

Db 1910 ATGCCCGCAGCGGAGATCTCGTGTGACCATGGCGATGCTGCTTGGCGAATATCATG 1969

QY 978 GTGGAATATGGCGCTTTTCTGGATTATCATGATGCGCGCTGGGTGGCGGACCGC 1037

Db 1970 GTGGAATATGGCGCTTTTCTGGATTATCATGATGCGCGCTGGGTGGCGGACCGC 2029

QY 1038 TATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGCGAATGGCT 1097

Db 2030 TATCAGGACATAGCGTTGGCTACCCGTGATATTGCTGAAGAGCTTGGCGCGAATGGCT 2089

QY 1098 GACCGCTTCTCGTCTTTACCGTATCGCGCTCCCGATTCGAGCGCATTCGCTTCTAT 1157

Db 2090 GACCGCTTCTCGTCTTTACCGTATTCGCGCTCCCGATTCGAGCGCATTCGCTTCTAT 2149

QY 1158 CGCTTCTTACGAGTTCTTCTGAGTTTAAACAGACCAACGTTTCCCTCTAGCGGGA 1217

Db 2150 CGCTTCTTACGAGTTCTTCTGAGTTTAAACAGACCAACGTTTCCCTCTAGCGGGA 2209

QY 1218 TCAATTCCGCCCTCTCCCTCCCCCCCCCTAACGTTACTGCCGAAGCGCTTGGAAATA 1277

Db 2210 TCAATTCCGCCCTCTCCCTCCCCCCCCCTAACGTTACTGCCGAAGCGCTTGGAAATA 2269

QY 1278 AGCGCGTGTGGTTTGTCTATATGTTATTTTCCACCATATTTGCGCTCTTTTGGCAATGT 1337

Db 2270 AGCGCGTGTGGTTTGTCTATATGTTATTTTCCACCATATTTGCGCTCTTTTGGCAATGT 2329

QY 1338 GAGGCGCCGAAACCTGGCCCTCTTCTTTCAGAGAGATTCCTAGGGCTTTTCCCTCT 1397

Db 2330 GAGGCGCCGAAACCTGGCCCTCTTCTTTCAGAGAGATTCCTAGGGCTTTTCCCTCT 2389

QY	1398	CGCCAAAGGAATGCMAGTCTGTGTAATGTCGTGAAGAGCAGTTCCTCTGGAGCTTC	1457	QY	2478	GCTTGTCTCTGAAACCGTCCTCGCGCCACCTAGGTTTCGGGGGCTATATCTCTAAGGC	2537
Db	2390	CGCCAAAGGAATGCAAGTCTGTGTAATGTCGTGAAGAGCAGTTCCTCTGGAGCTTC	2449	Db	3470	GCTTGTCTCTGAAACCGTCCTCGCGCCACCTAGGTTTCGGGGGCTATATCTCTAAGGC	3529
QY	1458	TTGAAGACAAACAAAGTCTGTAGCAGCCCTTTGACGAGCAGGGAACCCCACTTGGCGA	1517	QY	2538	ACATGGTATCGACCCCTAAACATCAGAACCGGGGTAAAGAACCATCACACGGGTGCCCCAT	2597
Db	2450	TTGAAGACAAACAAAGTCTGTAGCAGCCCTTTGACGAGCAGGGAACCCCACTTGGCGA	2509	Db	3530	ACATGGTATCGACCCCTAAACATCAGAACCGGGGTAAAGAACCATCACACGGGTGCCCCAT	3589
QY	1518	CAGTGTCTCTGGGCGCAAAAGCCACGCTGTATAGATACACTCTGCAAAAGGCGGCAACACC	1577	QY	2598	CACGTACTCTACCTATGCAAGTTCCTTTCGCGAGCGGTGCTCTCTGGGGGCGCTATGA	2657
Db	2510	CAGTGTCTCTGGGCGCAAAAGCCACGCTGTATAGATACACTCTGCAAAAGGCGGCAACACC	2569	Db	3590	CACGTACTCTACCTATGCAAGTTCCTTTCGCGAGCGGTGCTCTCTGGGGGCGCTATGA	3649
QY	1578	CCAGTGTCCACGTTGAGTTCGATAGTTCGGAAGAGTCAAAAGGCTCTCTCTCAAGCGT	1637	QY	2658	CATCATATATGTGATGAGTGCACCTCAACTCACTCGACCACTATCTCTGGGCATCGGCAC	2717
Db	2570	CCAGTGTCCACGTTGAGTTCGATAGTTCGGAAGAGTCAAAAGGCTCTCTCTCAAGCGT	2629	Db	3650	CATCATATATGTGATGAGTGCACCTCAACTCACTCGACCACTATCTCTGGGCATCGGCAC	3709
QY	1638	ATTCAACAAAGGGCTGAAGGATGCCCAAGAGTACCCCATTTGATGGGATCTGATCTGGG	1697	QY	2718	AGTCTGTGACCAAGCGGAGAGCGGTGGAGCGCACTCGTCTGCTCGCCACCGCTACGCC	2777
Db	2630	ATTCAACAAAGGGCTGAAGGATGCCCAAGAGTACCCCATTTGATGGGATCTGATCTGGG	2689	Db	3710	AGTCTGTGACCAAGCGGAGAGCGGTGGAGCGCACTCGTCTGCTCGCCACCGCTACGCC	3769
QY	1698	GCCTCGTGCACATGCTTTACATGTTTGTAGTTCGAGGTTAAAAAGCTTAGGCCCCCGG	1757	QY	2778	TCCGGGATCGCTCACCGTGCACATCCAAACATCGAGGAGTGGCTCTCTCCAGCACTGG	2837
Db	2690	GCCTCGTGCACATGCTTTACATGTTTGTAGTTCGAGGTTAAAAAGCTTAGGCCCCCGG	2749	Db	3770	TCCGGGATCGCTCACCGTGCACATCCAAACATCGAGGAGTGGCTCTCTCCAGCACTGG	3829
QY	1758	AACCAAGGGGAGCTGTTTCTCTTTGAAAAACACGATTAATACCATGGCGCTTATACGGC	1817	QY	2838	AGAAATCCCTTTTATGCAAAAGCCATCCCATCGAGACCATCAAGGGGGGAGGACCT	2897
Db	2750	AACCAAGGGGAGCTGTTTCTCTTTGAAAAACACGATTAATACCATGGCGCTTATACGGC	2809	Db	3830	AGAAATCCCTTTTATGCAAAAGCCATCCCATCGAGACCATCAAGGGGGGAGGACCT	3889
QY	1818	CTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCTAGCTTCACAGGCGGGGA	1877	QY	2898	CATTTTCTGCCATTCCAAAGAAATGTGATGAGTTCGCGGGAAGCTGTCTCGGCTCGG	2957
Db	2810	CTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCTAGCTTCACAGGCGGGGA	2869	Db	3890	CATTTTCTGCCATTCCAAAGAAATGTGATGAGTTCGCGGGAAGCTGTCTCGGCTCGG	3949
QY	1878	CAGGAACCAAGTTCGAGGGGAGGTCGAAGTGTCTCCACCGCAACACAAATCTTCTGGC	1937	QY	2958	ACTCAATCTCTAGCATATACCGGGGCTTGTATGTCATCCGTCATACCACTTAGCGGGA	3017
Db	2870	CAGGAACCAAGTTCGAGGGGAGGTCGAAGTGTCTCCACCGCAACACAAATCTTCTGGC	2929	Db	3950	ACTCAATCTCTAGCATATACCGGGGCTTGTATGTCATCCGTCATACCACTTAGCGGGA	4009
QY	1938	GACCTGCTCAATGAGCGTGTGTCGACTGTCTATCATGTGCGCGCTCAAGACCTTGC	1997	QY	3018	CGTCATCTCTGTAGCAACGCGCTCTAATGACGGGCTTTACCGGCGATTTTCGACTCAGT	3077
Db	2930	GACCTGCTCAATGAGCGTGTGTCGACTGTCTATCATGTGCGCGCTCAAGACCTTGC	2989	Db	4010	CGTCATCTCTGTAGCAACGCGCTCTAATGACGGGCTTTACCGGCGATTTTCGACTCAGT	4069
QY	1998	CGGCCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGACCAAGGACCTCGTGGGTG	2057	QY	3078	GATCGACTGCAATACATGTGTACCCAGACAGTTCGACTGAGCTTGAGCCCGACCTTCAC	3137
Db	2990	CGGCCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGACCAAGGACCTCGTGGGTG	3049	Db	4070	GATCGACTGCAATACATGTGTACCCAGACAGTTCGACTGAGCTTGAGCCCGACCTTCAC	4129
QY	2058	GCAAGGCCCCCGGGGCGTTCCTTGACACCATGCACTCGCGGAGCTCGGACCTTTA	2117	QY	3138	CATTCAGACGAGCCGTCGCCACAGACGCGGTGTACGCTCGAGCGCGGAGGAGGAC	3197
Db	3050	GCAAGGCCCCCGGGGCGTTCCTTGACACCATGCACTCGCGGAGCTCGGACCTTTA	3109	Db	4130	CATTCAGACGAGCCGTCGCCACAGACGCGGTGTACGCTCGAGCGCGGAGGAGGAC	4189
QY	2118	CTTGTGTACAGGCAATGCGATGTCTATTCGGGTGCGCGGGGCGACAGAGGGGAG	2177	QY	3198	TGGTAGGGGCGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGAGAACCGCCCTCGGGCAT	3257
Db	3110	CTTGTGTACAGGCAATGCGATGTCTATTCGGGTGCGCGGGGCGACAGAGGGGAG	3169	Db	4190	TGGTAGGGGCGAGGATGGGCAATTTACAGGTTTGTGACTCCAGGAGAACCGCCCTCGGGCAT	4249
QY	2178	CCTACTCTCCCCAGGCGCGTCTCTACTTGAAGGCTCTTCGGGCGGTCCACTGCTCTG	2237	QY	3258	GTTTCAATTCCTCGGTTCTGTGGAGTGTATGACGCGGGCTGTGTCTGTGTAGAGTCTAC	3317
Db	3170	CCTACTCTCCCCAGGCGCGTCTCTACTTGAAGGCTCTTCGGGCGGTCCACTGCTCTG	3229	Db	4250	GTTTCAATTCCTCGGTTCTGTGGAGTGTATGACGCGGGCTGTGTCTGTGTAGAGTCTAC	4309
QY	2238	CCCTCTGGGGCAGCTGTGGGATCTTTCGGGCTCGCTGTGCAACCGAGGGGTTGCGAA	2297	QY	3318	GCCTCCCGAGACCTCAGTTAGTTCGGGCTTACCTTAAACACACAGGGTTCGCCGTCTG	3377
Db	3230	CCCTCTGGGGCAGCTGTGGGATCTTTCGGGCTCGCTGTGCAACCGAGGGGTTGCGAA	3289	Db	4310	GCCTCCCGAGACCTCAGTTAGTTCGGGCTTACCTTAAACACACAGGGTTCGCCGTCTG	4369
QY	2298	GGCGGTGACATTTGTACCCGTGAGTCTATGAAACCACTATGCGGTCCCGGTCTTTCAC	2357	QY	3378	CCAGGACCATCTGGAGTTCGGGAGAGCTTTTACAGGCTTCACCCCATAGAGCCGCA	3437
Db	3290	GGCGGTGACATTTGTACCCGTGAGTCTATGAAACCACTATGCGGTCCCGGTCTTTCAC	3349	Db	4370	CCAGGACCATCTGGAGTTCGGGAGAGCTTTTACAGGCTTCACCCCATAGAGCCGCA	4429
QY	2358	GGACAACTGTCCTCCCGCGGTACCGCAGACATTCGAGGTGGGCCATCTACAGCCGC	2417	QY	3438	TTTCTTGTCCAGACTAAGCAGGAGGAGACACTTCCCTACCTCTGTGTAGCATACAGGC	3497
Db	3350	GGACAACTGTCCTCCCGCGGTACCGCAGACATTCGAGGTGGGCCATCTACAGCCGC	3409	Db	4430	TTTCTTGTCCAGACTAAGCAGGAGGAGACACTTCCCTACCTCTGTGTAGCATACAGGC	4489
QY	2418	TACTGTAGCGCAGAGACCTAAGTCCCGCTCGGTATGAGCCCAAGGATTAAGT	2477	QY	3498	TACGTTGTGCGCAGGCTCAGGCTCCACCTCCATCTGTGGGACCAAAATGTGAAGTGTCT	3557
Db	3410	TACTGTAGCGCAGAGACCTAAGTCCCGCTCGGTATGAGCCCAAGGATTAAGT	3469	Db	4490	TACGTTGTGCGCAGGCTCAGGCTCCACCTCCATCTGTGGGACCAAAATGTGAAGTGTCT	4549
				QY	3558	CATACGGGTAAAGCCTACGCTGCAAGGCGCCAAAGCCCCCTGTGTATAGGCTGGAGCCGT	3617

|||||  
4550 CATACGGCTAAGCCTACCGTGCACGGGCAACGCCCTCTGCTATAGGCTGGAGCCGT 4609  
QY  
3618 TCAAAACGAGGTACTACACACACACCCCAATACCAATACATCATGGCATGATGTCGC 3677  
Db  
4610 TCAAAACGAGGTACTACACACACACCCCAATACCAATACATCATGGCATGATGTCGC 4669  
QY  
3678 TGACCTGGAGGTGCTACAGACACCTGGGTGCTGGTAGCGGAGTCTTAGCAGCTCTGSC 3737  
Db  
4670 TGACCTGGAGGTGCTACAGACACCTGGGTGCTGGTAGCGGAGTCTTAGCAGCTCTGSC 4729  
QY  
3738 CGCGTATTCCTGACAAACAGGACGCTGCTCATTTGTGGCAGGATCATCTTGTCCGAAA 3797  
Db  
4730 CGCGTATTCCTGACAAACAGGACGCTGCTCATTTGTGGCAGGATCATCTTGTCCGAAA 4789  
QY  
3798 GCGGCGCATATTCGCGACAGGAAGTCTTTACCGGAGTTGATGAGATGGAAGAGTG 3857  
Db  
4790 GCGGCGCATATTCGCGACAGGAAGTCTTTACCGGAGTTGATGAGATGGAAGAGTG 4849  
QY  
3858 CGCCTCACACCTCCCTTATCATCGAAACAGGAATGAGCTCGCGGAAACATTCAAACAGAA 3917  
Db  
4850 CGCCTCACACCTCCCTTATCATCGAAACAGGAATGAGCTCGCGGAAACATTCAAACAGAA 4909  
QY  
3918 GSCAATCGGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCTGCTCCCGTGTGGA 3977  
Db  
4910 GSCAATCGGGTTGCTGCAAAACAGCCACCAAGCAAGCGAGGCTGCTCCCGTGTGGA 4969  
QY  
3978 ATCCAAAGTGGCGGACCTCGAAGCCTTCTGGCGAAGCATATGTGGAATTTATCAGCGG 4037  
Db  
4970 ATCCAAAGTGGCGGACCTCGAAGCCTTCTGGCGAAGCATATGTGGAATTTATCAGCGG 5029  
QY  
4038 GATACAAATTTAGCAGGCTTGTCCACTGTGCTGGCAACCGCGGATGACATCTGAT 4097  
Db  
5030 GATACAAATTTAGCAGGCTTGTCCACTGTGCTGGCAACCGCGGATGACATCTGAT 5089  
QY  
4098 GGCATTACAGCCTCTATCACAGCCGCTCACCAACCAACATACCCCTCTGTTTAAACAT 4157  
Db  
5090 GGCATTACAGCCTCTATCACAGCCGCTCACCAACCAACATACCCCTCTGTTTAAACAT 5149  
QY  
4158 CTGGGGGATGGTGGCGGCCCACTTGTCTCTCCAGCGTGTCTGCTTTCGTAGG 4217  
Db  
5150 CTGGGGGATGGTGGCGGCCCACTTGTCTCTCCAGCGTGTCTGCTTTCGTAGG 5209  
QY  
4218 CCGCGCATCGTGGAGCGGCTGTTGGACGATAGCGCTTGGAGGCTGTTGGATAT 4277  
Db  
5210 CCGCGCATCGTGGAGCGGCTGTTGGACGATAGCGCTTGGAGGCTGTTGGATAT 5269  
QY  
4278 TTTGGCAGGTTATGGAGCAGGGGTGGCAGCGGCTCGTGGCTTTAAGTCTATGAGCGG 4337  
Db  
5270 TTTGGCAGGTTATGGAGCAGGGGTGGCAGCGGCTCGTGGCTTTAAGTCTATGAGCGG 5329  
QY  
4338 CGAGATGCCCTCACCGAGGACCTGGTTAACTACTCTCCCTGCTATCTCTCCCTGGCG 4397  
Db  
5330 CGAGATGCCCTCACCGAGGACCTGGTTAACTACTCTCCCTGCTATCTCTCCCTGGCG 5389  
QY  
4398 CTTAGTCTGGGGTCTGTCGACGCGATATCTGCTCGCACGCTGGGCGCAGGGAGGG 4457  
Db  
5390 CTTAGTCTGGGGTCTGTCGACGCGATATCTGCTCGCACGCTGGGCGCAGGGAGGG 5449  
QY  
4458 GGCCTGTGAGTGGATGAACCGGCTGATAGGTTTCGCTTCGCGGGTAAACACGCTCTCC 4517  
Db  
5450 GGCCTGTGAGTGGATGAACCGGCTGATAGGTTTCGCTTCGCGGGTAAACACGCTCTCC 5509  
QY  
4518 CACGCACTATGTGCTGAGAGGACGCTGCAGCAGTGTCACTCAGATCCTCTTAGTCT 4577  
Db  
5510 CACGCACTATGTGCTGAGAGGACGCTGCAGCAGTGTCACTCAGATCCTCTTAGTCT 5569  
QY  
4578 TACCATCACTAGCTGCTGAAGGGCTTCAACAGTGGATCAAGGAGTCTGTCCACGCC 4637  
Db  
5570 TACCATCACTAGCTGCTGAAGGGCTTCAACAGTGGATCAAGGAGTCTGTCCACGCC 5629  
QY  
4638 ATGCTCCGCTCGTGGCTAAGAGATCTTTGGGATTCGATATGACCGGTGTTGACTGATTT 4697  
|||||

Db  
5630 ATGCTCCGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGACCGGTGTTGACTGATTT 5689  
QY  
4698 CAAGACCTGGCTCCAGTCCAAGCTCCTGCCGGAATTTGCCGGAGTCCCTTCTTCTCATG 4757  
Db  
5690 CAAGACCTGGCTCCAGTCCAAGCTCCTGCCGGAATTTGCCGGAGTCCCTTCTTCTCATG 5749  
QY  
4758 TCAAGTGGGTACAGGGAGTCTGCGGGGCGACGGCATCATGCAAAACACCTGCCCATG 4817  
Db  
5750 TCAAGTGGGTACAGGGAGTCTGCGGGGCGACGGCATCATGCAAAACACCTGCCCATG 5809  
QY  
4818 TGGAGCACAGATCACCGGACATGTGAAAACCGGTTCCATGAGATCGTGGGGCTTAGGAC 4877  
Db  
5810 TGGAGCACAGATCACCGGACATGTGAAAACCGGTTCCATGAGATCGTGGGGCTTAGGAC 5869  
QY  
4878 CTGTAGTAAACAGTGGCATGGAACATTTCCCAATTAAACGCTACACACCGGGCCCTGCAC 4937  
Db  
5870 CTGTAGTAAACAGTGGCATGGAACATTTCCCAATTAAACGCTACACACCGGGCCCTGCAC 5929  
QY  
4938 GCGCTCCCGGGCGCCAAATTTCTAGGCGCTGTGGCGGTGGCTGCTGAGAGTACGT 4997  
Db  
5930 GCGCTCCCGGGCGCCAAATTTCTAGGCGCTGTGGCGGTGGCTGCTGAGAGTACGT 5989  
QY  
4998 GGAGGTTACCGGGTGGGGATTTCCACTACGCTGACCGGCATGACCACTGCAACGTAAA 5057  
Db  
5990 GGAGGTTACCGGGTGGGGATTTCCACTACGCTGACCGGCATGACCACTGCAACGTAAA 6049  
QY  
5058 GTGCCGTGTGAGTTCCGGCCCCCGAAATTTCTTCAACAGAGTGGATGGGTGCGTTCGA 5117  
Db  
6050 GTGCCGTGTGAGTTCCGGCCCCCGAAATTTCTTCAACAGAGTGGATGGGTGCGTTCGA 6109  
QY  
5118 CAGGTACGCTCAGCGTGCMAACCCCTCTACGGGAGGATGACATTCCTGTCGGGCT 5177  
Db  
6110 CAGGTACGCTCAGCGTGCMAACCCCTCTACGGGAGGATGACATTCCTGTCGGGCT 6169  
QY  
5178 CAATCAATACCTGTTGGGTACAGCTCCCATGCCAGCCGACCGAGTACGAGTGT 5237  
Db  
6170 CAATCAATACCTGTTGGGTACAGCTCCCATGCCAGCCGACCGAGTACGAGTGT 5229  
QY  
5238 CACTTCCATGCTACCGACCCCTCCCAATTAACGGCGGAGACGGCTAAGCGTAGGCTGC 5297  
Db  
6230 CACTTCCATGCTACCGACCCCTCCCAATTAACGGCGGAGACGGCTAAGCGTAGGCTGC 6289  
QY  
5298 CAGGGATCTCCCGCTCTTGGCAGCTCATCAGTACGACGCTGTCTGCGGCTTCTT 5357  
Db  
6290 CAGGGATCTCCCGCTCTTGGCAGCTCATCAGTACGACGCTGTCTGCGGCTTCTT 6349  
QY  
5358 GAAGGCAACATCATACCCGCTCATGCTCCCGGACGCTGACCTCATCGAGCCAACT 5417  
Db  
6350 GAAGGCAACATCATACCCGCTCATGCTCCCGGACGCTGACCTCATCGAGCCAACT 6409  
QY  
5418 CTTGTGGCGCAGGAGATGGCGGGAACATACCCCGTGGAGTCAGAAAAATAAGGTAGT 5477  
Db  
6410 CTTGTGGCGCAGGAGATGGCGGGAACATACCCCGTGGAGTCAGAAAAATAAGGTAGT 6469  
QY  
5478 AATTTTGGACTCTTTTCGACCGCTCCAGAGGAGGATGAGAGGATATCCGTTCC 5537  
Db  
6470 AATTTTGGACTCTTTTCGACCGCTCCAGAGGAGGATGAGAGGATATCCGTTCC 6529  
QY  
5538 GCGGAGATCTCTCGGAGGTCCAGAAATTTCCCTCGAGGATGCCCATATGGGCAACGCC 5597  
Db  
6530 GCGGAGATCTCTCGGAGGTCCAGAAATTTCCCTCGAGGATGCCCATATGGGCAACGCC 6589  
QY  
5598 GGATTAACACCTTCACTGTTAGAGTCTTGAAGAACCCCGGACTACGTCCTCCAGTGT 5657  
Db  
6590 GGATTAACACCTTCACTGTTAGAGTCTTGAAGAACCCCGGACTACGTCCTCCAGTGT 6649  
QY  
5658 ACACGGTGTCCATTCGCGCTGCGCAAGGCCCTCCGATACCACTCCACGAGGAGAG 5717  
Db  
6650 ACACGGTGTCCATTCGCGCTGCGCAAGGCCCTCCGATACCACTCCACGAGGAGAG 6709  
QY  
5718 GACGGTGTCTGTGAGAAATCTACCGTGTCTTCTGCTGGCGGAGTCTGCCACAAAGAC 5777  
Db  
6710 GACGGTGTCTGTGAGAAATCTACCGTGTCTTCTGCTGGCGGAGTCTGCCACAAAGAC 6769

QY 5778 CTTGGGAGCTCCGATCGTGGCGCTGAGCAGCGGACGCGAAGCGCTCTCTTGACCA 5837  
Db 6770 CTTGGGAGCTCCGATCGTGGCGCTGAGCAGCGGACGCGAAGCGCTCTCTTGACCA 6829  
QY 5838 GCCCTCCGACGACGCGGAGCGGGATCCGATCGTGGAGTGGTACTCTCTCATGCCCCCCT 5897  
Db 6810 GCCCTCCGACGACGCGGAGCGGGATCCGATCGTGGAGTGGTACTCTCTCATGCCCCCCT 6889  
QY 5898 TGAGGGGAGCGGGGGATCCGATCTCAGCGAGGGTCTTGGTCTACCGTAAGCGAGGA 5957  
Db 6890 TGAGGGGAGCGGGGGATCCGATCTCAGCGAGGGTCTTGGTCTACCGTAAGCGAGGA 6949  
QY 5958 GGCTAGTCAGGAGCTGCTGCTGCTGATCTCTACATCTGATGAGTGGAGCGGCGCTGATCAC 6017  
Db 6950 GGCTAGTCAGGAGCTGCTGCTGCTGATCTCTACATCTGATGAGTGGAGCGGCGCTGATCAC 7009  
QY 6018 GCCATGCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTGAGCAACTCTTTGCTCG 6077  
Db 7010 GCCATGCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTGAGCAACTCTTTGCTCG 7069  
QY 6078 TCACCAAACTTGTCTATGCTAACAATCTCGAGCGCAAGCTGCGGCGAGAAAGGT 6137  
Db 7070 TCACCAAACTTGTCTATGCTAACAATCTCGAGCGCAAGCTGCGGCGAGAAAGGT 7129  
QY 6138 CACCTTTGACAGACTGCAAGTCTGAGCGACCACTACCGGACGCTGCTCAAGGAGATGAA 6197  
Db 7130 CACCTTTGACAGACTGCAAGTCTGAGCGACCACTACCGGACGCTGCTCAAGGAGATGAA 7189  
QY 6198 GCGGAAGCGTCCAGTAAAGGTAAACTTCTATCCGTGGAGGAAGCTGTAAAGCTGAC 6257  
Db 7190 GCGGAAGCGTCCAGTAAAGGTAAACTTCTATCCGTGGAGGAAGCTGTAAAGCTGAC 7249  
QY 6258 GCCCCCACTTGGCGAGATCTAAATTTGGCTATGGGCGAAGGACGCTCGGAACTTATC 6317  
Db 7250 GCCCCCACTTGGCGAGATCTAAATTTGGCTATGGGCGAAGGACGCTCGGAACTTATC 7309  
QY 6318 CAGCAAGCGGTAAACCATCTCGCTCGTGGAGGACTTGTGGAGGACACTGAGAC 6377  
Db 7310 CAGCAAGCGGTAAACCATCTCGCTCGTGGAGGACTTGTGGAGGACACTGAGAC 7369  
QY 6378 ACCAATTGACACCACTCATGCGCAAAATAGAGTTTCTCGCTCCCAACAGAGAGGG 6437  
Db 7370 ACCAATTGACACCACTCATGCGCAAAATAGAGTTTCTCGCTCCCAACAGAGAGGG 7429  
QY 6438 GGGCGCAAGCGAGTCCCTTATGCTATCCAGATTTGGGGTTGCTGTGGAGAA 6497  
Db 7430 GGGCGCAAGCGAGTCCCTTATGCTATCCAGATTTGGGGTTGCTGTGGAGAA 7489  
QY 6498 AATGGCCCTTTACGATGCTGCTCACCTCCCTCAGGCGCTGATGGCTCTTCATACGG 6557  
Db 7490 AATGGCCCTTTACGATGCTGCTCACCTCCCTCAGGCGCTGATGGCTCTTCATACGG 7549  
QY 6558 ATTCCAATCTCTCTGACAGCGGTCGAGTTCTGGTGAATGCTCGGAAGCGAAGAA 6617  
Db 7550 ATTCCAATCTCTCTGACAGCGGTCGAGTTCTGGTGAATGCTCGGAAGCGAAGAA 7609  
QY 6618 ATGCGCTATGGCTTCGATATGACACCGCTGTTTGAATCAACGCTCACTGAGATGA 6677  
Db 7610 ATGCGCTATGGCTTCGATATGACACCGCTGTTTGAATCAACGCTCACTGAGATGA 7669  
QY 6678 CATCGCTTGGAGGTCAATCTACCAATGTTGACTTGGCTCCCGGAGCGACAGAGG 6737  
Db 7670 CATCGCTTGGAGGTCAATCTACCAATGTTGACTTGGCTCCCGGAGCGACAGAGG 7729  
QY 6738 CATAAGTGGCTTCAGAGCGGCTTTACATCGGGGCCCCCTGACTAATTTAAAGGCA 6797  
Db 7730 CATAAGTGGCTTCAGAGCGGCTTTACATCGGGGCCCCCTGACTAATTTAAAGGCA 7789  
QY 6798 GAACTGCGGCTATCGCGGTGCGCGAGGCGGTGCTACTGACGACCGCTGCGGTAAATAC 6857  
Db 7790 GAACTGCGGCTATCGCGGTGCGCGAGGCGGTGCTACTGACGACCGCTGCGGTAAATAC 7849

QY 6858 CCTCACATGTTACTTTGAAGCGCGCTGCGGCTGTGAGCTGCGAAGCTCCAGGACTGCAC 6917  
Db 7850 CCTCACATGTTACTTTGAAGCGCGCTGCGGCTGTGAGCTGCGAAGCTCCAGGACTGCAC 7909  
QY 6918 GATGCTCGTATGCGGAGACGACCTTTGTCGTTATCTGTGAAAGCGCGGAGCCCAAGAGGA 6977  
Db 7910 GATGCTCGTATGCGGAGACGACCTTTGTCGTTATCTGTGAAAGCGCGGAGCCCAAGAGGA 7969  
QY 6978 CGAGGCGAGCTACGCGGCTTACGAGGCTATGACTAGATATCTGTCGCCGCCCTCGGGA 7037  
Db 7970 CGAGGCGAGCTACGCGGCTTACGAGGCTATGACTAGATATCTGTCGCCGCCCTCGGGA 8029  
QY 7038 CCGGCCAAACAGAAATACGACTTTGGAGTTGATTAACATCATGCTCTCTCAATGTCAGT 7097  
Db 8030 CCGGCCAAACAGAAATACGACTTTGGAGTTGATTAACATCATGCTCTCTCAATGTCAGT 8089  
QY 7098 CGGCGAGATGATCTGGCAAAAGGGTGTACTTCTCAACCGTGAACCCACACCCCT 7157  
Db 8090 CGGCGAGATGATCTGGCAAAAGGGTGTACTTCTCAACCGTGAACCCACACCCCT 8149  
QY 7158 TGCGGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAATTCCTGGCTAGGCAACAT 7217  
Db 8150 TGCGGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAATTCCTGGCTAGGCAACAT 8209  
QY 7218 CATCATGATGCGGCCACTTTGTGGCAAGGATGATCCTGATGACTCATTTCTTCTCCAT 7277  
Db 8210 CATCATGATGCGGCCACTTTGTGGCAAGGATGATCCTGATGACTCATTTCTTCTCCAT 8269  
QY 7278 CTTTCTAGCTCAGGAACAACTTGAAAAAGCCCTAGATTGTGAGATCTACGGGGCTGTGA 7337  
Db 8270 CTTTCTAGCTCAGGAACAACTTGAAAAAGCCCTAGATTGTGAGATCTACGGGGCTGTGA 8329  
QY 7338 CTCCATTGAGCCACTTGACCTACCTCAGATCATTTCAAGCACTCCATGCGCTTAGGCAAT 7397  
Db 8330 CTCCATTGAGCCACTTGACCTACCTCAGATCATTTCAAGCACTCCATGCGCTTAGGCAAT 8389  
QY 7398 TTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCACTCCCTCAGAAACT 7457  
Db 8390 TTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCACTCCCTCAGAAACT 8449  
QY 7458 TGGGTTACCGCCCTTGCGAGTCTGGAGACATCGGGCCAGAAAGTGTCCGCGTAGGCTACT 7517  
Db 8450 TGGGTTACCGCCCTTGCGAGTCTGGAGACATCGGGCCAGAAAGTGTCCGCGTAGGCTACT 8509  
QY 7518 GTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTCACTCTTCACTGGGCGAGTAAGAC 7577  
Db 8510 GTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTCACTCTTCACTGGGCGAGTAAGAC 8569  
QY 7578 CAAGCTCAAACTCACTCCAATCCCGCTGCGTCCAGTTGGATTATCCAGCTGGTTCGT 7637  
Db 8570 CAAGCTCAAACTCACTCCAATCCCGCTGCGTCCAGTTGGATTATCCAGCTGGTTCGT 8629  
QY 7638 TGCTGGTTACAGCGGGGAGACATATATCAAGCCTGTCTCGTCCCGACCCCGCTGGTT 7697  
Db 8630 TGCTGGTTACAGCGGGGAGACATATATCAAGCCTGTCTCGTCCCGACCCCGCTGGTT 8689  
QY 7698 CATGTGTGCTACTCTCTACTTCTGTAGGGTAGGCACTATCTACTCCCGACCCGATG 7757  
Db 8690 CATGTGTGCTACTCTCTACTTCTGTAGGGTAGGCACTATCTACTCCCGACCCGATG 8749  
QY 7758 AACGGGACCTTAAACACTCCAGGCAATAGGCCATCTCTGTTTTTTTTCCTTTTTTTTTT 7817  
Db 8750 AACGGGACCTTAAACACTCCAGGCAATAGGCCATCTCTGTTTTTTTTCCTTTTTTTTTT 8809  
QY 7818 CTTT 7877  
Db 8810 CTTT 8869  
QY 7878 CTTTTTTTCTTCTTGGTGGCTCCATCTTAGCCCTAGTCAACGCTAGCTGTGAAAGGCTCG 7937  
Db 8870 CTTTTTTCTTCTTGGTGGCTCCATCTTAGCCCTAGTCAACGCTAGCTGTGAAAGGCTCG 8929  
QY 7938 TGAGCCGCTTGACTGCGAGAGAGTGTGATCTGCGCTCTCTGCGATCAAGT 7989



```
Db      8930  TGAGCCGCTTGACTGCAGAGAGTGCTGATGACTGAGCTCTCTGCGCTCTCTGAGATCAAGT 8981
|||||
RESULT 15
AX937621
LOCUS      AX937621                12315 bp      DNA      linear      PAT 06-JAN-2004
DEFINITION Sequence 1 from Patent WO03091439.
ACCESSION  AX937621
VERSION     AX937621.1  GI:40713672
KEYWORDS
SOURCE      unidentified
ORGANISM    unidentified
REFERENCE   1
  AUTHORS   Duggal,R.A., Patrick,A.K., Zhang,J.A. and Zhao,W.A.
  TITLE     Reporter-selectable hepatitis c virus replicon
  JOURNAL   Patent: WO 03091439-A 1 06-NOV-2003;
            PFIZER INC. (US)
FEATURES
  source
    1..12315
    /organism="unidentified"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32644"
    /note="A cell line wherein the nucleic acid molecule is a
    self replicating RNA molecule."
ORIGIN
Query Match      95.2%; Score 7607.2; DB 6; Length 12315;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7609; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 378  GGGCGCGCATGATTGAAACAAGATGATTGCACGAGTTCTCCGGCCGCTTGGGTGGAG 437
Db 1540  GGGCGCGCATGATTGAAACAAGATGATTGCACGAGTTCTCCGGCCGCTTGGGTGGAG 1599
QY 438  AGCGTATTCGGCTATGACTGGGACACAGACAAATCGGCTGCTCGATCGCCCGGTTC 497
Db 1600  AGCGTATTCGGCTATGACTGGGACACAGACAAATCGGCTGCTCGATCGCCCGGTTC 1659
QY 498  CGGCTGTACAGCGAGGGCGCCCGGTTCCTTTTGTCAAGACCGACCTGTCGGGTGCCCTG 557
Db 1660  CGGCTGTACAGCGAGGGCGCCCGGTTCCTTTTGTCAAGACCGACCTGTCGGGTGCCCTG 1719
QY 558  AATGAATCGAGACGAGCGAGCGCGGTATCGTGGCTGCCACGACGCGCGTTCCTTGC 617
Db 1720  AATGAATCGAGACGAGCGAGCGCGGTATCGTGGCTGCCACGACGCGCGTTCCTTGC 1779
QY 618  GCAGCTGTCTCGAGTTCCTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGGAAGTG 677
Db 1780  GCAGCTGTCTCGAGTTCCTCACTGAAGCGGGAAGGACTGGCTGCTATTGGCGGAAGTG 1839
QY 678  CGGGGCGAGGATCTCCTGTATCTCACTTGTCTTGTCCGCGAAGAACTATCCATCATGGCT 737
Db 1840  CGGGGCGAGGATCTCCTGTATCTCACTTGTCTTGTCCGCGAAGAACTATCCATCATGGCT 1899
QY 738  GATCAATCGCGGGCTGCATACGCTTGATCGGCTACTCTGCCATTCGACCAACCAAGCG 797
Db 1900  GATCAATCGCGGGCTGCATACGCTTGATCGGCTTACCTGCCATTCGACCAACCAAGCG 1959
QY 798  AAACATCGCATCGAGCGACGATCTCGATCGGATGGAGCCGGTCTTGTCCATCAGGATGAT 857
Db 1960  AAACATCGCATCGAGCGACGATCTCGGATGGAGCCGGTCTTGTCCATCAGGATGAT 2019
QY 858  CTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGGAATCTTCCGCAAGGCTCAAGGCGCGC 917
Db 2020  CTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGGAATCTTCCGCAAGGCTCAAGGCGCGC 2079
QY 918  ATGCCGACGCGGAGGATCTCGTGTGACCCATCGCGATGCCGTGCTGCCGAATATCATG 977
Db 2080  ATGCCGACGCGGAGGATCTCGTGTGACCCATCGCGATGCCGTGCTGCCGAATATCATG 2139
QY 978  GTGGAATATGGCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGGCGGACCGC 1037
|||||
Db      8930  TGAGAAAATGGCCGCTTTTCTGATTCATCGACTGTGGCCGCTGGGTGGCGGACCGC 2199
1038  TATCAGGACATAGCGTTGGCTACCGTGATATTCCTGAAGAGCTTGGCGGCGGAATGGCT 1097
2200  TATCAGGACATAGCGTTGGCTACCGTGATATTCCTGAAGAGCTTGGCGGCGGAATGGCT 2259
1098  GACCGCTTCTCGTGCTTTACCGGTATCGCGCTTCCGATTCGAGCGCATTCGCTTCTAT 1157
2260  GACCGCTTCTCGTGCTTTACCGGTATCGCGCTTCCGATTCGAGCGCATTCGCTTCTAT 2319
1158  CGCGTCTTGGAGAGTCTTCTGAGTTTAAACAGACCAACAGGTTTCCCTCTAGCGGA 1217
2320  CGCGTCTTGGAGAGTCTTCTGAGTTTAAACAGACCAACAGGTTTCCCTCTAGCGGA 2379
1218  TCAATTCGCGCCCTCTCCCTCCCGCCCTTAACTTCTGCGCCGAGCGCTTGGAAATA 1277
2380  TCAATTCGCGCCCTCTCCCTCCCGCCCTTAACTTCTGCGCCGAGCGCTTGGAAATA 2439
1278  AGGCGGTGTGGTGTCTATATGTTATTTCCACATATTTGCGCTTTTGGCAATGT 1337
2440  AGGCGGTGTGGTGTCTATATGTTATTTCCACATATTTGCGCTTTTGGCAATGT 2499
1338  GAGGCGCGGAAACCTGGCCCTGCTCTTGTGACGACATTTCTAGGGTCTTTCCCTCT 1397
2500  GAGGCGCGGAAACCTGGCCCTGCTCTTGTGACGACATTTCTAGGGTCTTTCCCTCT 2559
1398  CGCCAAAGGAATGCAAGGTCTCTTGAATTCGTGAAGAGAGTTCCTCTGGAAGCTTC 1457
2560  CGCCAAAGGAATGCAAGGTCTCTTGAATTCGTGAAGAGAGTTCCTCTGGAAGCTTC 2619
1458  TTGAAGACAAACAAACGCTGTGTAGCGACCTTTTGAGGCGAGCGGAAACCCCTC 1517
2620  TTGAAGACAAACAAACGCTGTGTAGCGACCTTTTGAGGCGAGCGGAAACCCCTC 2679
1518  CAGGTGCTCTCGGCGCAAAAGCCAGCTGTATAAGATACCTGGAAGCGCGCAACCC 1577
2680  CAGGTGCTCTCGGCGCAAAAGCCAGCTGTATAAGATACCTGGAAGCGCGCAACCC 2739
1578  CCAGTGCCACAGTGTGAGTTGATAGTTGTGAAAGAGTCAAAATGGCTCTCTCAAGCGT 1637
2740  CCAGTGCCACAGTGTGAGTTGATAGTTGTGAAAGAGTCAAAATGGCTCTCTCAAGCGT 2799
1638  ATTCAACAAAGGGCTGAAGGATGCCAGAAAGTACCCCATTTGTATGGATCTGATCGG 1697
2800  ATTCAACAAAGGGCTGAAGGATGCCAGAAAGTACCCCATTTGTATGGATCTGATCGG 2859
1698  GCCTCGGTGCATGCTTTACATGTTTGTAGTTCGAGGTTTAAAGAGCTTAGGCCCGCG 1757
2860  GCCTCGGTGCATGCTTTACATGTTTGTAGTTCGAGGTTTAAAGAGCTTAGGCCCGCG 2919
1758  AACCACGGGCGAGTGTGTTTCTTTTGAAGAAACAGATATACCATGGCGCTTATACGCG 1817
2920  AACCACGGGCGAGTGTGTTTCTTTTGAAGAAACAGATATACCATGGCGCTTATACGCG 2979
1818  CTACTCCCAACAGACGCGAGGCTCTATGGCTGTGATCATCATAGCTCAGAGCGCGGA 1877
2980  CTACTCCCAACAGACGCGAGGCTCTATGGCTGTGATCATCATAGCTCAGAGCGCGGA 3039
1878  CAGGAACAGAGTTCGAGGGGAGTCCAGTGGTCTCCACCGCAACACATCTTTCTGGC 1937
3040  CAGGAACAGAGTTCGAGGGGAGTCCAGTGGTCTCCACCGCAACACATCTTTCTGGC 3099
1938  GACCTGCGTCAATGGCGTGTGTTGACTGTCTATCATGGTCCGCTCAAGACCTTGC 1997
3100  GACCTGCGTCAATGGCGTGTGTTGACTGTCTATCATGGTCCGCTCAAGACCTTGC 3159
1998  CGGCGCAAGGGCGCAATCACCAGATGTACCAATGTGACCAAGGACCTCGTGGGTG 2057
3160  CGGCGCAAGGGCGCAATCACCAGATGTACCAATGTGACCAAGGACCTCGTGGGTG 3219
2058  GCAAGCGCCCGCGGGCGCTTCTTGTACACCATGCACTCGCGAGCTCGGACCTTTA 2117
|||||
```

Db 3220 GAAAGCGCCCGCGGCGCTCTCTTGACACCATGACCTCGGCGAGCTCGACCTTTA 3279  
QY 2118 CTTGGTACAGGAGATGCGGATGTCATTTCCGGTGGCGCGGCGGCGGAG 2177  
Db 3280 CTTGGTACAGGAGATGCGGATGTCATTTCCGGTGGCGCGGCGGCGGAG 3339  
QY 2178 CCTACTCTCCCGCGAGCGCGCTCTCTACTTGAAGGGCTCTTCGGGCGGTCACTGCTCTG 2237  
Db 3340 CCTACTCTCCCGCGAGCGCGCTCTCTACTTGAAGGGCTCTTCGGGCGGTCACTGCTCTG 3399  
QY 2238 CCCCTCGGGGACCGGTGGGATCTTTTCGGGCTGCGGTGTCACCCGAGGGGTTGCGAA 2297  
Db 3400 CCCCTCGGGGACCGGTGGGATCTTTTCGGGCTGCGGTGTCACCCGAGGGTTGCGAA 3459  
QY 2298 GCGGGTGGACTTTGTACCGGTGAGTCTATGGAACCACTATGCGGTCCCGGTCTTCAC 2357  
Db 3460 GCGGGTGGACTTTGTACCGGTGAGTCTATGGAACCACTATGCGGTCCCGGTCTTCAC 3519  
QY 2358 GGAACAATCGTCCCTCGCGCGGTACCGAGACATTCAGGTGGGCCATCTACACGCCCC 2417  
Db 3520 GGAACAATCGTCCCTCGCGCGGTACCGAGACATTCAGGTGGGCCATCTACACGCCCC 3579  
QY 2418 TACTGGTAGCGCAAGACAATAAGTCCCGGTGCGTATGACGCCCAAGGTTAAGGT 2477  
Db 3580 TACTGGTAGCGCAAGACAATAAGTCCCGGTGCGTATGACGCCCAAGGTTAAGGT 3639  
QY 2478 GCTTGCTGAAACCGGTGCGCGCGCACCTAGTTTCGGGGGTATATGTTCTAAGGC 2537  
Db 3640 GCTTGCTGAAACCGGTGCGCGCGCACCTAGTTTCGGGGGTATATGTTCTAAGGC 3699  
QY 2538 ACATGGTATCGACCTTAACATCAGAACCGGGTAAGACCACTACACCGGGTGCCCCAT 2597  
Db 3700 ACATGGTATCGACCTTAACATCAGAACCGGGTAAGACCACTACACCGGGTGCCCCAT 3759  
QY 2598 CACGTACTCCACTATGGCAAGTTCTTTCGCGACGGTGGTGTCTCTGGGGGCGCTATGA 2657  
Db 3760 CACGTACTCCACTATGGCAAGTTCTTTCGCGACGGTGGTGTCTCTGGGGGCGCTATGA 3819  
QY 2658 CATCATATATGATGATGTCACCTCACTGACTCGACCACTATCTTGGGCGATCGGCAC 2717  
Db 3820 CATCATATATGATGATGTCACCTCACTGACTCGACCACTATCTTGGGCGATCGGCAC 3879  
QY 2718 AGTCTGACCAAGCGGAGAGCGGTGGAGCGGACTCGTCTGCTCGCCACCGCTACGCC 2777  
Db 3880 AGTCTGACCAAGCGGAGAGCGGTGGAGCGGACTCGTCTGCTCGCCACCGCTACGCC 3939  
QY 2778 TCCGGGATCGGTACCGGTGCCATATCCAAACATCGAGAGGTGGTCTGTCCAGACTGG 2837  
Db 3940 TCCGGGATCGGTACCGGTGCCATATCCAAACATCGAGAGGTGGTCTGTCCAGACTGG 3999  
QY 2838 AGAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATCAAGGGGGGAGGCACCT 2897  
Db 4000 AGAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATCAAGGGGGGAGGCACCT 4059  
QY 2898 CATTTCTGCGCATTCGAAGAAATGTGATGAGCTCGCGCGAAGCTGTCCGGGCTCGG 2957  
Db 4060 CATTTCTGCGCATTCGAAGAAATGTGATGAGCTCGCGCGAAGCTGTCCGGGCTCGG 4119  
QY 2958 ACTCAATGCTGTAGCATATACCGGGGCTGTATGTATCCGTATACCAACTAGCGGAGA 3017  
Db 4120 ACTCAATGCTGTAGCATATACCGGGGCTGTATGTATCCGTATACCAACTAGCGGAGA 4179  
QY 3018 CGTCACTGCTGTAGCAACGGCGCTCTAAAGCGGGCTTTACCGGCGATTTTCGACTCAGT 3077  
Db 4180 CGTCACTGCTGTAGCAACGGCGCTCTAAAGCGGGCTTTACCGGCGATTTTCGACTCAGT 4239  
QY 3078 GATCGACTGCAATACATGTGTACCCAGACAGTCTGACTTTCAGCCCTGGACCCGACCTTCAC 3137  
Db 4240 GATCGACTGCAATACATGTGTACCCAGACAGTCTGACTTTCAGCCCTGGACCCGACCTTCAC 4299  
QY 3138 CATTCAGACGACCGGTGCCAAGACGCGGTGTACGCTCGACGCGGCGAGGAGGAC 3197  
Db 4300 CATTCAGACGACCGGTGCCAAGACGCGGTGTACGCTCGACGCGGCGAGGAGGAC 4359

QY 3198 TGGTAGGGCAGGATGGGCAATTTACAGGTTGTGACTCCAGGAGAACGCGCTTCGGGCAT 3257  
Db 4360 TGGTAGGGCAGGATGGGCAATTTACAGGTTGTGACTCCAGGAGAACGCGCTTCGGGCAT 4419  
QY 3258 GTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGTGCTTGATACGACTCAC 3317  
Db 4420 GTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGTGCTTGATACGACTCAC 4479  
QY 3318 GCCCGCGAGACCTCAGTTAGGTTGCGGGCTTAACTAAACAACACAGAGGTTGCCGCTGTG 3377  
Db 4480 GCCCGCGAGACCTCAGTTAGGTTGCGGGCTTAACTAAACAACACAGAGGTTGCCGCTGTG 4539  
QY 3378 CCAGGACCATCTGAGGTTCTGGGAGAGCGTCTTTACAGGCGCTCACCCACATAGAGCCCA 3437  
Db 4540 CCAGGACCATCTGAGGTTCTGGGAGAGCGTCTTTACAGGCGCTCACCCACATAGAGCCCA 4599  
QY 3438 TTTCTTGTCACAGACTAAGCAGGAGGAGACAACTTCCCTTACCTGGTAGCATACAGGC 3497  
Db 4600 TTTCTTGTCACAGACTAAGCAGGAGGAGACAACTTCCCTTACCTGGTAGCATACAGGC 4659  
QY 3498 TACGGTGTGCGCAGGGCTCAGGCTCACCCTCCATCTGTTGGACCCAAATGTGGAAGTGTCT 3557  
Db 4660 TACGGTGTGCGCAGGGCTCAGGCTCACCCTCCATCTGTTGGACCCAAATGTGGAAGTGTCT 4719  
QY 3558 CATACGGCTAAGGCTACGCTGCACGCGCCAAACGCGCTGTGTATAGCTGGGAGCGCT 3617  
Db 4720 CATACGGCTAAGGCTACGCTGCACGCGCCAAACGCGCTGTGTATAGCTGGGAGCGCT 4779  
QY 3618 TCAAAACAGAGTTACTACACACACACCCCATTAACAAATACATCATGCGATCATGTCGCG 3677  
Db 4780 TCAAAACAGAGTTACTACACACACACCCCATTAACAAATACATCATGCGATCATGTCGCG 4839  
QY 3678 TGACCTGAGAGTGTCTCAGGACACCTGGTGTCTGTGTTAGGCGGAGTCTTAGCAGCTCTGCG 3737  
Db 4840 TGACCTGAGAGTGTCTCAGGACACCTGGTGTCTGTGTTAGGCGGAGTCTTAGCAGCTCTGCG 4899  
QY 3738 CGCGTATTGCTGACACAGCAGCGTGTCTATGTTGGGCGAGGATCATCTTGTCCGGA 3797  
Db 4900 CGCGTATTGCTGACACAGCAGCGTGTCTATGTTGGGCGAGGATCATCTTGTCCGGA 4959  
QY 3798 GCCGCGCATCATCTCCGACAGGGAAGTCTTTTACCGGGAGTTCGATGAGATGGAGAGTG 3857  
Db 4960 GCCGCGCATCATCTCCGACAGGGAAGTCTTTTACCGGGAGTTCGATGAGATGGAGAGTG 5019  
QY 3858 CGCTCACAACCTCCCTTACATCGAACAGGGAATGAGCTCCCGGAACAAATTCAAACAGAA 3917  
Db 5020 CGCTCACAACCTCCCTTACATCGAACAGGGAATGAGCTCCCGGAACAAATTCAAACAGAA 5079  
QY 3918 GGCAATCGGTTGTCGAAACAGCCACCAAGCAGGAGGCTGCTGCTCCGCTGGTGA 3977  
Db 5080 GGCAATCGGTTGTCGAAACAGCCACCAAGCAGGAGGCTGCTGCTCCGCTGGTGA 5139  
QY 3978 ATCCAAGTGGCGGACCTTCGAAGCTTCTGGGCGAAGCATATGTGGAATTTTCATCAGCGG 4037  
Db 5140 ATCCAAGTGGCGGACCTTCGAAGCTTCTGGGCGAAGCATATGTGGAATTTTCATCAGCGG 5199  
QY 4038 GATCAATATTTTAGAGGCTTGTGACTCTGCTGGGGAACCCCGGATAGCATCACTGAT 4097  
Db 5200 GATCAATATTTTAGAGGCTTGTGACTCTGCTGGGGAACCCCGGATAGCATCACTGAT 5259  
QY 4098 GGCAATCAGGCTTCTATCAGCGCGCTCACACCCACATACCCCTCTCTGTTTAAACAT 4157  
Db 5260 GGCAATCAGGCTTCTATCAGCGCGCTCACACCCACATACCCCTCTCTGTTTAAACAT 5319  
QY 4158 CCTGGGGGATGGGTGGCGCCAACTTGTCTCTCCAGCGCTGCTTCTGTTTCTGATAGG 4217  
Db 5320 CCTGGGGGATGGGTGGCGCCAACTTGTCTCTCCAGCGCTGCTTCTGTTTCTGATAGG 5379  
QY 4218 CGCGGCTATCGCTGGAGCGGCTGTGGCAGCATAGGCTTGGGAGGTGCTTGTGCTGATAT 4277  
Db 5380 CGCGGCTATCGCTGGAGCGGCTGTGGCAGCATAGGCTTGGGAGGTGCTTGTGCTGATAT 5439

QY 4278 TTTGGCAGGTTATGGACAGGGGTGGCAGGCGCGCTCGTGCCCTTTAAGGTCATGACGG 4337  
Db 5440 TTTGGCAGGTTATGGACAGGGGTGGCAGGCGCGCTCGTGCCCTTTAAGGTCATGACGG 5499  
QY 4338 CGAGATGCCCTTCCACCGAGGACCTGGTTAACTACTCCTCTCTATCTCTCCCTGGCG 4397  
Db 5500 CGAGATGCCCTTCCACCGAGGACCTGGTTAACTACTCCTCTCTATCTCTCCCTGGCG 5559  
QY 4398 CCTAGTCGTGGGGTCTGTGCGCAGCGATACCTGCTGCGCACTGTGGGCCACGAGGGAGG 4457  
Db 5560 CCTAGTCGTGGGGTCTGTGCGCAGCGATACCTGCTGCGCACTGTGGGCCACGAGGGAGG 5619  
QY 4458 GGGTGTGCACTGAGTAACCGGCTGATAGCTTTCGCTTCGGGGTAAACACGTCCTCCC 4517  
Db 5620 GGGTGTGCACTGAGTAACCGGCTGATAGCTTTCGCTTCGGGGTAAACACGTCCTCCC 5679  
QY 4518 CAGCACTATGTGCTGAGAGCGACGCTGAGCAGCTGTCACTCAGATCCTCTCTAGTCT 4577  
Db 5680 CAGCACTATGTGCTGAGAGCGACGCTGAGCAGCTGTCACTCAGATCCTCTCTAGTCT 5739  
QY 4578 TACCATCACTCAGCTGCTGAGAGGCTTACCAGTGGATCAACGAGACTGCTCCAGCC 4637  
Db 5740 TACCATCACTCAGCTGCTGAGAGGCTTACCAGTGGATCAACGAGACTGCTCCAGCC 5799  
QY 4638 ATGCTCGGCTCGTGGCTAAGAGATGTTGGGATTGGATATGACGCTGTGACTGATTT 4697  
Db 5800 ATGCTCGGCTCGTGGCTAAGAGATGTTGGGATTGGATATGACGCTGTGACTGATTT 5859  
QY 4698 CAAGACTGCTCAGTCCAAGCTCTCTGCGCGATTCGCGGAGTCCCTTCTCTCATG 4757  
Db 5860 CAAGACTGCTCAGTCCAAGCTCTCTGCGCGATTCGCGGAGTCCCTTCTCTCATG 5919  
QY 4758 TCAACGCTGGCTCAAGGAGCTCTGGCGGGGAGCGGATCATGCAACACCTGCCCAGT 4817  
Db 5920 TCAACGCTGGCTCAAGGAGCTCTGGCGGGGAGCGGATCATGCAACACCTGCCCAGT 5979  
QY 4818 TGGAGCACAGATCACCGGACATGTGAAAAACGGTTCCATGAGGATCGTGGGCTTAGGAC 4877  
Db 5980 TGGAGCACAGATCACCGGACATGTGAAAAACGGTTCCATGAGGATCGTGGGCTTAGGAC 6039  
QY 4878 CTGTAGTAACAGTGGCATGGAACATTTCCCATTTAACGCGTACACACGCGGCCCTGCAC 4937  
Db 6040 CTGTAGTAACAGTGGCATGGAACATTTCCCATTTAACGCGTACACACGCGGCCCTGCAC 6099  
QY 4938 GGCCTCCCGGCGCCAAATTTATCTAGGCGCTGTGGCGGGTGGCTGCTGAGGAGTACGT 4997  
Db 6100 GGCCTCCCGGCGCCAAATTTATCTAGGCGCTGTGGCGGGTGGCTGCTGAGGAGTACGT 6159  
QY 4998 GGAGGTTACGCGGGTGGGGATTTCCACTACGTGACGGGATGACCACTGACACACGTAA 5057  
Db 6160 GGAGGTTACGCGGGTGGGGATTTCCACTACGTGACGGGATGACCACTGACACACGTAA 6219  
QY 5058 GTGCCCGTGTAGGTTCCGGCCCGGAAATTTTACAGAGTGGATGGGGTGGGTTGCA 5117  
Db 6220 GTGCCCGTGTAGGTTCCGGCCCGGAAATTTTACAGAGTGGATGGGGTGGGTTGCA 6279  
QY 5118 CAGGTACGCTCCAGCTGCAACCCCTCTACCGGAGAGGTCACTTCTGTGCGGCT 5177  
Db 6280 CAGGTACGCTCCAGCTGCAACCCCTCTACCGGAGAGGTCACTTCTGTGCGGCT 6339  
QY 5178 CAATCAATACCTGGTGGGTACAGCTCCCATCGAGCCCGCAACCGAGCTAGCAGTCT 5237  
Db 6340 CAATCAATACCTGGTGGGTACAGCTCCCATCGAGCCCGCAACCGAGCTAGCAGTCT 6399  
QY 5238 CACTTCCATGCTCACGACCCCTCCCAATTCGCGGAGAGCGGCTAAGCGTAGGCTGGC 5297  
Db 6400 CACTTCCATGCTCACGACCCCTCCCAATTCGCGGAGAGCGGCTAAGCGTAGGCTGGC 6459  
QY 5298 CAGGGATCTCCCCCTCTCTGGCAGCTCATCAGCTAGCAGCTGTCTGGCCTTCTT 5357  
Db 6460 CAGGGATCTCCCCCTCTCTGGCAGCTCATCAGCTAGCAGCTGTCTGGCCTTCTT 6519  
QY 5358 GAAGGCAACATGCACTACCGCTCATGACTCCCGGACGCTGACCTCATCGAGCCCACT 5417

Db 6520 GAAGGCAACATGCACTACCGCTCATGACTCCCGGACGCTGACCTCATCGAGGCCAACCT 6579  
QY 5418 CCTGTGCGCGCAGGAGATGGCGGAAATATCACCCGCTGGAGTCAGAAATAGGTAGT 5477  
Db 6580 CCTGTGCGCGCAGGAGATGGCGGAAATATCACCCGCTGGAGTCAGAAATAGGTAGT 6639  
QY 5478 AATTTTGGACTTTTCGAGCGCTTCCAAAGCGGAGGAGATGAGAGGAAATATCCGTTCC 5537  
Db 6640 AATTTTGGACTTTTCGAGCGCTTCCAAAGCGGAGGAGATGAGAGGAAATATCCGTTCC 6699  
QY 5538 GCGGAGATCCTGCGGAGTCCAGGAAATTCCTCGAGCGATGCCATATGCGGACGCCC 5597  
Db 6700 GCGGAGATCCTGCGGAGTCCAGGAAATTCCTCGAGCGATGCCATATGCGGACGCCC 6759  
QY 5598 GGATTACAAACCTTCCACTGTAGAGTCTCGAGGACCCGAGCTACGCTCCCTCAGTGGT 5657  
Db 6760 GGATTACAAACCTTCCACTGTAGAGTCTCGAGGACCCGAGCTACGCTCCCTCAGTGGT 6819  
QY 5658 ACACGGGTGCCATTTGCCGCTGCCAAGGCCCTCCGATACCACTCCACGGAGGAAGAG 5717  
Db 6820 ACACGGGTGCCATTTGCCGCTGCCAAGGCCCTCCGATACCACTCCACGGAGGAAGAG 6879  
QY 5718 GACGGTGTCTCTCAGAACTACCGTGTCTTCTGCTTGGCGAGCTGCCCAAGAGAC 5777  
Db 6880 GACGGTGTCTCTCAGAACTACCGTGTCTTCTGCTTGGCGAGCTGCCCAAGAGAC 6939  
QY 5778 CTTTCGGCAGCTCCGAATTCGTCGGCGCTGACAGCGCACGCGAACGCGCTCTCTGACCA 5837  
Db 6940 CTTTCGGCAGCTCCGAATTCGTCGGCGCTGACAGCGCACGCGAACGCGCTCTCTGACCA 6999  
QY 5838 GCCTTCGAGCAGCGGAGCGGAGTCCGACGTTGAGTCTCTCTCTCCATGCCGCCCT 5897  
Db 7000 GCCTTCGAGCAGCGGAGCGGAGTCCGACGTTGAGTCTCTCTCTCCATGCCGCCCT 7059  
QY 5898 TGAGGGGAGCGCGGGATCCCGATCTCAGCAGCGGTCTTGGTCTACCGTAAGCAGGA 5957  
Db 7060 TGAGGGGAGCGCGGGATCCCGATCTCAGCAGCGGTCTTGGTCTACCGTAAGCAGGA 7119  
QY 5958 GGCTAGTAGGAGCGTCTGCTGCTCGATGTCTACATGACAGCGGCCCTGATCAC 6017  
Db 7120 GGCTAGTAGGAGCGTCTGCTGCTCGATGTCTACATGACAGCGGCCCTGATCAC 7179  
QY 6018 GCCATCGCTGGAGGAGGAAACCAAGCTGCCCATCAATGCACTGAGCAACTCTTTGCTCCG 6077  
Db 7180 GCCATCGCTGGAGGAGGAAACCAAGCTGCCCATCAATGCACTGAGCAACTCTTTGCTCCG 7239  
QY 6078 TCACCAACATTTGGTCTATGCTACACATCTCGCAGCGAGCTCGGGCAGAGAGGT 6137  
Db 7240 TCACCAACATTTGGTCTATGCTACACATCTCGCAGCGAGCTCGGGCAGAGAGGT 7299  
QY 6138 CACTTTGACAGACTCGAGTCTGAGACCACTACCGGAGCTCTCAAGGAGATGA 6197  
Db 7300 CACTTTGACAGACTCGAGTCTGAGACCACTACCGGAGCTCTCAAGGAGATGA 7359  
QY 6198 GCGGAAGGCTCCAGTTAAGCTAAACTCTATCCGTGGAGGAGCTGTAAAGTAC 6257  
Db 7360 GCGGAAGGCTCCAGTTAAGCTAAACTCTATCCGTGGAGGAGCTGTAAAGTAC 7419  
QY 6258 GCCCCACATTCGCGAGATCTAAATTTGGCTATGGGCAAGAGCTCGGACCTATC 6317  
Db 7420 GCCCCACATTCGCGAGATCTAAATTTGGCTATGGGCAAGAGCTCGGACCTATC 7479  
QY 6318 CAGCAAGGCGGTTAACCACTACCGCTCCGTGTTGAGGAGCTTGTGGAAGACACTGAGAC 6377  
Db 7480 CAGCAAGGCGGTTAACCACTACCGCTCCGTGTTGAGGAGCTTGTGGAAGACACTGAGAC 7539  
QY 6378 ACCAATTGACACCAATCATGGCAAAATAGGTTTCTCGTCCCAACCGAGAGGG 6437  
Db 7540 ACCAATTGACACCAATCATGGCAAAATAGGTTTCTCGTCCCAACCGAGAGGG 7599  
QY 6438 GGGCGCAGCCAGCTCGCTTATCGTATCCAGATTTGGGGGTCTGTGTGCGAGAA 6497

Db 7600 GGGCGGCAAGCAGCTCGCTTATCGTATCCAGATTTGGGGGTTCTGTTGTGCGAGAA 7659  
 Qy 6498 AATGGCCCTTTACGATGTTGTTCTCCACCTCCCTCAGGCGGTGATGGGCTCTTCTATACGG 6557  
 Db 7660 AATGGCCCTTTACGATGTTGTTCTCCACCTCCCTCAGGCGGTGATGGGCTCTTCTATACGG 7719  
 Qy 6558 ATTCCCATACTCTCTCGACAGCGGTCGAGTCTCTGGTGAATGCTCGAAGCGAAGAA 6617  
 Db 7720 ATTCCCATACTCTCTCGACAGCGGTCGAGTCTCTGGTGAATGCTCGAAGCGAAGAA 7779  
 Qy 6618 ATGCCCTATGGGCTTCGATATGACACCGGCTGTTTTCAGCTCAACGGTCACTGAGAATGA 6677  
 Db 7780 ATGCCCTATGGGCTTCGATATGACACCGGCTGTTTTCAGCTCAACGGTCACTGAGAATGA 7839  
 Qy 6678 CATCGGTTGAGGAGTCAATCTACCAATGTTGTGACTTGGCCCCCGGAAGCGACAGAGC 6737  
 Db 7840 CATCGGTTGAGGAGTCAATCTACCAATGTTGTGACTTGGCCCCCGGAAGCGACAGAGC 7899  
 Qy 6738 CATAAGGTCGCTCAAGAGCGGCTTTACATCGGGGGCCCCCTGACTAATTTCTAAAGGGCA 6797  
 Db 7900 CATAAGGTCGCTCAAGAGCGGCTTTACATCGGGGGCCCCCTGACTAATTTCTAAAGGGCA 7959  
 Qy 6798 GAACTGCGGCTATCGCGGTCGCGCGAGCGGTGTACTGACGACAGCTCGGTAAATAC 6857  
 Db 7960 GAACTGCGGCTATCGCGGTCGCGCGAGCGGTGTACTGACGACAGCTCGGTAAATAC 8019  
 Qy 6858 CCTACATGTTACTTTGAAGCGCGCTGCGGCTGTGAGTGGGAAGCTCCAGGACTGCAC 6917  
 Db 8020 CCTACATGTTACTTTGAAGCGCGCTGCGGCTGTGAGTGGGAAGCTCCAGGACTGCAC 8079  
 Qy 6918 GATGCTCTATCGGAGACGACTTGTGTTATCTGTGAAGCGGGGACCCCAAGAGCA 6977  
 Db 8080 GATGCTCTATCGGAGACGACTTGTGTTATCTGTGAAGCGGGGACCCCAAGAGCA 8139  
 Qy 6978 CGAGCGAGCCTACCGGCGCTTACGAGGCTATGACTAGATCTCTGCCCGCCCTGGGGA 7037  
 Db 8140 CGAGCGAGCCTACCGGCGCTTACGAGGCTATGACTAGATCTCTGCCCGCCCTGGGGA 8199  
 Qy 7038 CCGGCGGCAACGAGATAGGCTTGGAGTTGATATCATCATGCTCTCTCCATGTTGCTAGT 7097  
 Db 8200 CCGGCGGCAACGAGATAGGCTTGGAGTTGATATCATCATGCTCTCTCCATGTTGCTAGT 8259  
 Qy 7098 CGCGCAGCATGCTATCGGCAAGGCTGACTATCTCACCGGTGACCCACACCCCGCT 7157  
 Db 8260 CGCGCAGCATGCTATCGGCAAGGCTGACTATCTCACCGGTGACCCACACCCCGCT 8319  
 Qy 7158 TCGCGGGCTGCGTGGGACAGCTAGACACACTCCAGTCAATCTCTGGCTAGGCAACAT 7217  
 Db 8320 TCGCGGGCTGCGTGGGACAGCTAGACACACTCCAGTCAATCTCTGGCTAGGCAACAT 8379  
 Qy 7218 CATCATGATGCGGCGCCTGTGGCAAGGATGATCTGATGACTCATTTCTTCTCCAT 7277  
 Db 8360 CATCATGATGCGGCGCCTGTGGCAAGGATGATCTGATGACTCATTTCTTCTCCAT 8439  
 Qy 7278 CTTTCTAGCTCAGGAACAACCTTGAAGGCGCTAGATTCTCAGATCTACGGGCGCTGTTA 7337  
 Db 8440 CTTTCTAGCTCAGGAACAACCTTGAAGGCGCTAGATTCTCAGATCTACGGGCGCTGTTA 8499  
 Qy 7338 CTCCATGAGCCACTTGACCTACCTCAGATCAATCAAGACTCCATGGCCTTAGGCAATT 7397  
 Db 8500 CTCCATGAGCCACTTGACCTACCTCAGATCAATCAAGACTCCATGGCCTTAGGCAATT 8559  
 Qy 7398 TTCACTCCATAGTTACTCTCCAGGTGAGTCAATAGGTTGGCTTTCATGCTCAGGAACCT 7457  
 Db 8560 TTCACTCCATAGTTACTCTCCAGGTGAGTCAATAGGTTGGCTTTCATGCTCAGGAACCT 8619  
 Qy 7458 TGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCGAAGTGTCCGGCTAGGCTACT 7517  
 Db 8620 TGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCGAAGTGTCCGGCTAGGCTACT 8679  
 Qy 7518 GTCCCAAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTTCACTGGGCGAGTAAGGAC 7577  
 Db 8680 GTCCCAAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTTCACTGGGCGAGTAAGGAC 8739

Qy 7578 CAAGCTCAAACTCACTCCAAATCCGGCTGCGTCCAGTTGGATTATCCAGCTGGTTCGT 7637  
 Db 8740 CAAGCTCAAACTCACTCCAAATCCGGCTGCGTCCAGTTGGATTATCCAGCTGGTTCGT 8799  
 Qy 7638 TGCTGGTTACAGCGGGGAGACATATATCACAGCCTGTCTCGTGGCCGACCCCGTGGTT 7697  
 Db 8800 TGCTGGTTACAGCGGGGAGACATATATCACAGCCTGTCTCGTGGCCGACCCCGTGGTT 8859  
 Qy 7698 CATGTGGTGCCTACTCCTACTCTTCTGTAGGGGTAGGCAATCTATCTACTCCCAACCGATG 7757  
 Db 8860 CATGTGGTGCCTACTCCTACTCTTCTGTAGGGGTAGGCAATCTATCTACTCCCAACCGATG 8919  
 Qy 7758 AACGGGACCTAAACACTCCAGGCGCAATAGGCCATCTCTGTTTTCCTCTTTTTTTTTT 7817  
 Db 8920 AACGGGACCTAAACACTCCAGGCGCAATAGGCCATCTCTGTTTTCCTCTTTTTTTTTT 8979  
 Qy 7818 CTTC 7877  
 Db 8980 CTTC 9039  
 Qy 7878 CTTTTCTTTCCCTTTGGTGGCTCCATCTTAGCCCTAGTCACGGCTAGCTGTGAAGSTCCG 7937  
 Db 9040 CTTTTCTTTCCCTTTGGTGGCTCCATCTTAGCCCTAGTCACGGCTAGCTGTGAAGSTCCG 9099  
 Qy 7938 TGAGCCGCTTGACTGCAGAGAGTGTCTGATCTGGGCTCTCTGCAGATCAAGT 7989  
 Db 9100 TGAGCCGCTTGACTGCAGAGAGTGTCTGATCTGGGCTCTCTGCAGATCAAGT 9151

Search completed: December 9, 2004, 12:31:02  
 Job time : 32078 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 18:04:05 ; Search time 3311 Seconds

(without alignments)  
12666.151 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989

Sequence: 1 gccagcccccgcattgggggc.....ggcctctctgcagatcaagt 7989

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 413486 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N Geneseq\_23Sep04:\*  
1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7989	100.0	7989	6	AAD25322
2	7989	100.0	11313	12	ADP86264
3	7987.4	100.0	7989	3	AAA98968
4	7987.4	100.0	7989	6	AAD25326
5	7987.4	100.0	7989	12	ADJ57845
6	7987.4	100.0	7992	6	AAL47276
7	7987.4	100.0	10690	6	ABK91412
8	7987.4	100.0	10690	9	ACA61697
9	7987.4	100.0	10690	10	ADC83762
10	7987.4	100.0	11313	12	ADP86271
11	7985.8	100.0	7992	6	AAL47281
12	7985.8	100.0	10690	6	ABK91448
13	7985.8	100.0	10690	6	ABK91435
14	7985.8	100.0	10690	6	ABK91243
15	7985.8	100.0	10690	6	ABK91434
16	7985.8	100.0	11313	12	ADP86265
17	7984.2	99.9	7989	6	AAD25325
18	7984.2	99.9	7989	10	ADD93734
19	7984.2	99.9	10690	6	ABK91440
20	7984.2	99.9	11313	12	ADP86272
21	7982.6	99.9	7992	6	AAL47280

22	7982.6	99.9	7992	6	AAL47277
23	7981	99.9	11313	12	ADP86275
24	7979.4	99.9	10690	6	ABK91242
25	7979.4	99.9	11313	12	ADP86273
26	7977.8	99.9	11313	12	ADP86266
27	7976.4	99.8	10691	6	ABK91423
28	7976.2	99.8	11313	12	ADP86268
29	7974.4	99.8	10693	6	ABK91438
30	7973.4	99.8	7987	6	AAD25321
31	7973	99.8	11313	12	ADP86267
32	7972.8	99.8	10693	6	ABK91443
33	7971.8	99.8	7987	6	AAD25329
34	7971.6	99.8	10691	6	ABK91439
35	7970.2	99.8	7987	6	AAD25324
36	7969	99.7	7991	6	AAL47279
37	7965.4	99.7	8001	3	AAA98967
38	7961.6	99.7	7995	6	AAL47278
39	7957	99.6	11313	12	ADP86270
40	7953.8	99.6	11313	12	ADP86269
41	7779.4	97.4	7989	10	ADD93733
42	7713	96.5	11184	12	ADP86274
43	7705	96.4	11184	12	ADP86276
44	7700.2	96.4	11184	12	ADP86277
45	7697	96.3	7848	6	AAD25323

## ALIGNMENTS

### RESULT 1

AAD25322

ID AAD25322 standard; cDNA; 7989 BP.

XX AC

XX AAD25322;

DT 12-MAR-2002 (first entry)

XX Hepatitis C virus (HCV) repibBartMan/AvaiI cDNA.

XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;

KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;

KW ss.

XX Hepatitis C virus.

XX Key

FT CDS

FT Location/Qualifiers

FT 1801..7758

FT /tag= a

FT /product= "HCVrepibBartMan polyprotein"

FT misc\_feature

FT 7766

FT /tag= b

FT /note= "Nucleotide creating AvaiI site"

FT XX

XX WO200109364-A2.

PN XX

XX 29-NOV-2001.

PD XX

XX 23-MAY-2001; 2001WO-US016822.

PF XX

XX 23-MAY-2000; 2000US-00576989.

PR XX

XX (UNIW ) UNIV WASHINGTON.

PA XX

XX Rice CM, Blight KJ;

PI XX

XX WPI: 2002-066755/09.

DR XX

DR P-PSDB; AAE15717.

XX Hepatitis C virus variants having greater transfection efficiency and

PT ability to survive subpassage, useful as a vaccine for immunizing primate

PT to the virus, comprise non-naturally occurring viral sequences.

XX PS

PS Claim 44; Page 69-71; 174pp; English.





1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTGTTAGTCAGAGTTAAAA 1740 QY  
1681 ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTGTTAGTCAGAGTTAAAA 1740 Db  
1741 AAGCTCTAGGCCCCCGGAACCAAGGGGACCTGGTTTTCCTTTGAAAAACACGATAATACC 1800 QY  
1741 AAGCTCTAGGCCCCCGGAACCAAGGGGACCTGGTTTTCCTTTGAAAAACACGATAATACC 1800 Db  
1801 ATGGCGCTATTACGGCTACTCCCAACAGACGAGGCTACTTTGGCTGCATCATCACT 1860 QY  
1801 ATGGCGCTATTACGGCTACTCCCAACAGACGAGGCTACTTTGGCTGCATCATCACT 1860 Db  
1861 AGCTCACAGGCGGACACAGAACAGTGCAGAGGGAGTCCAAAGTGTCTCCACCGCA 1920 QY  
1861 AGCTCACAGGCGGACACAGAACAGTGCAGAGGGAGTCCAAAGTGTCTCCACCGCA 1920 Db  
1921 ACACAATCTTCTGGCGACCTCGTCAATGGCGTGTGTGGACTGTCTATCATGTGGCC 1980 QY  
1921 ACACAATCTTCTGGCGACCTCGTCAATGGCGTGTGTGGACTGTCTATCATGTGGCC 1980 Db  
1981 GGCTCAAGACCTTTCGCGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040 QY  
1981 GGCTCAAGACCTTTCGCGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040 Db  
2041 CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGCGCGTTTCTTGACACCATGCCTGC 2100 QY  
2041 CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGCGCGTTTCTTGACACCATGCCTGC 2100 Db  
2101 GGAGCTCGGACCTTTACTTGTGTGTCAGAGGATGCGGATGTCATTCGGGTGCGCGCGG 2160 QY  
2101 GGAGCTCGGACCTTTACTTGTGTGTCAGAGGATGCGGATGTCATTCGGGTGCGCGCGG 2160 Db  
2161 GGCGACAGCAGGGGAGCTACTCTCCCGAGCGCCGTCCTTACTTGAAGGCTCTTCG 2220 QY  
2161 GGCGACAGCAGGGGAGCTACTCTCCCGAGCGCCGTCCTTACTTGAAGGCTCTTCG 2220 Db  
2221 GCGGTCCTACTGCTCGCCCTCGGGGACGCTGTGGGCACTCTTCGGGCTGCGGTGTC 2280 QY  
2221 GCGGTCCTACTGCTCGCCCTCGGGGACGCTGTGGGCACTCTTCGGGCTGCGGTGTC 2280 Db  
2281 ACCCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGTCAATGTGAAAGGCTCTTCG 2340 QY  
2281 ACCCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGTCAATGTGAAAGGCTCTTCG 2340 Db  
2341 CGGTCCCCGGTCTTACGAGCAACTCGTCCCTCGCGCGGTACCGCAGACATTTCCAGGTG 2400 QY  
2341 CGGTCCCCGGTCTTACGAGCAACTCGTCCCTCGCGCGGTACCGCAGACATTTCCAGGTG 2400 Db  
2401 GCCCATCTACCGCCCTACTGCTAGCGGCAAGACACTAAGGTGCGCGCTGGTATGCA 2460 QY  
2401 GCCCATCTACCGCCCTACTGCTAGCGGCAAGACACTAAGGTGCGCGCTGGTATGCA 2460 Db  
2461 GCCCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGCGCCACCCCTAGGTTTCGGG 2520 QY  
2461 GCCCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGCGCCACCCCTAGGTTTCGGG 2520 Db  
2521 GGGTATGTCTTAAGGCACATGATCGACCTTAACATCAGAACCGGGGTAAAGACCATC 2580 QY  
2521 GGGTATGTCTTAAGGCACATGATCGACCTTAACATCAGAACCGGGGTAAAGACCATC 2580 Db  
2581 ACCAAGGGTGCCTTACGCTACTCCACTATGCAAGTCTTCTGCGACCGGTGGTTC 2640 QY  
2581 ACCAAGGGTGCCTTACGCTACTCCACTATGCAAGTCTTCTGCGACCGGTGGTTC 2640 Db  
2641 TCTGGGGCGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700 QY  
2641 TCTGGGGCGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700 Db  
2701 ATCTTGGGATCGGCACAGTCTTGGACCAAGCGGAGACGGCTGGAGCGCACTCGTGGT 2760 QY  
2701 ATCTTGGGATCGGCACAGTCTTGGACCAAGCGGAGACGGCTGGAGCGCACTCGTGGT 2760 Db

2761 CTGCCACCCCTAGGCTCCGGGATCGGTCAACCGTGCCACATCOAAACATCGAGAGGTG 2820 QY  
2761 CTGCCACCCCTAGGCTCCGGGATCGGTCAACCGTGCCACATCOAAACATCGAGAGGTG 2820 Db  
2821 GCTCTGTCCAGCACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880 QY  
2821 GCTCTGTCCAGCACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880 Db  
2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCCTCAAGAAAGAAATGTGATGAGTCCGCGC 2940 QY  
2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCCTCAAGAAAGAAATGTGATGAGTCCGCGC 2940 Db  
2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATAATACCGGGGCCCTTGTATTCGGTC 3000 QY  
2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATAATACCGGGGCCCTTGTATTCGGTC 3000 Db  
3001 ATACCAACTAGCGGACAGCTATTGTGTAGCAACCGAGCGCTCTAATGACGGGCTTTACC 3060 QY  
3001 ATACCAACTAGCGGACAGCTATTGTGTAGCAACCGAGCGCTCTAATGACGGGCTTTACC 3060 Db  
3061 GCGATTTTCGACTCAGTCACTGATCGACTCAATGTGTAGCAACCGAGCGCTCTAATGACGGGCTTTACC 3120 QY  
3061 GCGATTTTCGACTCAGTCACTGATCGACTCAATGTGTAGCAACCGAGCGCTCTAATGACGGGCTTTACC 3120 Db  
3121 CTGGACCCGACCTTACCAATGTAGACGACGACCGTGCACAGACGCGGTGTCAACGCTG 3180 QY  
3121 CTGGACCCGACCTTACCAATGTAGACGACGACCGTGCACAGACGCGGTGTCAACGCTG 3180 Db  
3181 CAGCGCGGAGGAGGACTGTGTAGGGCAGGATGGGCACTTTACAGGTTTGTGATCCAGGA 3240 QY  
3181 CAGCGCGGAGGAGGACTGTGTAGGGCAGGATGGGCACTTTACAGGTTTGTGATCCAGGA 3240 Db  
3241 GAAACGCGCTCGGGCACTTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300 QY  
3241 GAAACGCGCTCGGGCACTTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300 Db  
3301 GCTTGTAGAGTCAAGGCTCAGCGCCCGGAGACCTCAGTTAGGTTGCGGCTTACCTAAACACA 3360 QY  
3301 GCTTGTAGAGTCAAGGCTCAGCGCCCGGAGACCTCAGTTAGGTTGCGGCTTACCTAAACACA 3360 Db  
3361 CAGAGGTTCCCGCTGCGCAGGACCATCTGAGGTTCTGGAGAGCGTCTTTACAGGCTTC 3420 QY  
3361 CAGAGGTTCCCGCTGCGCAGGACCATCTGAGGTTCTGGAGAGCGTCTTTACAGGCTTC 3420 Db  
3421 ACCACATAGACGCCCATTTCTTGCCACAGACTAAGCAGGACGAGAGCAACTTCCCTTAC 3480 QY  
3421 ACCACATAGACGCCCATTTCTTGCCACAGACTAAGCAGGACGAGAGCAACTTCCCTTAC 3480 Db  
3481 CTGTTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCATCGTGGGAC 3540 QY  
3481 CTGTTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCATCGTGGGAC 3540 Db  
3541 CAAATGTGGAAGTGTCTCATACGCTTAAAGCCTTAAAGCCTTAAAGCCTTAAAGCCTTAAAGCCTT 3600 QY  
3541 CAAATGTGGAAGTGTCTCATACGCTTAAAGCCTTAAAGCCTTAAAGCCTTAAAGCCTTAAAGCCTT 3600 Db  
3601 TATAGGCTGGGAGCGCTTCAAAACGAGGTTACTACACACACCCCATAAACAAATACATC 3660 QY  
3601 TATAGGCTGGGAGCGCTTCAAAACGAGGTTACTACACACACCCCATAAACAAATACATC 3660 Db  
3661 ATGGATGATGTTCGGCTGACCTGAGGTCGTACAGGACCTGGGTGCTGTAGCGGA 3720 QY  
3661 ATGGATGATGTTCGGCTGACCTGAGGTCGTACAGGACCTGGGTGCTGTAGCGGA 3720 Db  
3721 GTCTAGCAGTCTGCGCGCGTATTTGCTGACACAGGAGGAGTCTTTACCGGAGTTC 3780 QY  
3721 GTCTAGCAGTCTGCGCGCGTATTTGCTGACACAGGAGGAGTCTTTACCGGAGTTC 3780 Db  
3781 ATCATCTTGTCCGGAAGCGGCGCATCATTTCCCGACAGGAGTCTTTACCGGAGTTC 3840 QY  
3781 ATCATCTTGTCCGGAAGCGGCGCATCATTTCCCGACAGGAGTCTTTACCGGAGTTC 3840 Db  
3841 GATGAGTGGAGAGTGGCGCTTACACCTCCCTTACATCGAACAGGGAATGCAGCTCGCC 3900 QY



QY	6061	AGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAAATCTCGAGCGCAAGC	6120
Db	6061	AGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAAATCTCGAGCGCAAGC	6120
QY	6121	CTGGCGAGAGAGAGTCACTTTTGACAGACTGCGAGTCTCTGGACGACCACTACCGGGAC	6180
Db	6121	CTGGCGAGAGAGAGTCACTTTTGACAGACTGCGAGTCTCTGGACGACCACTACCGGGAC	6180
QY	6181	GTGCTCAGGAGATGAAGGCGAAGCGTCCACAGTTAAGCTAACTTCTATCCGTGGAG	6240
Db	6181	GTGCTCAGGAGATGAAGGCGAAGCGTCCACAGTTAAGCTAACTTCTATCCGTGGAG	6240
QY	6241	GAAGCTGTAAAGCTGACGCGCCCACTTTCCGCCAGATCTAAATTTGGCTATGGGGCAAG	6300
Db	6241	GAAGCTGTAAAGCTGACGCGCCCACTTTCCGCCAGATCTAAATTTGGCTATGGGGCAAG	6300
QY	6301	GAGCTCCGGAACCTATCCAGCAAGCGCGTTAAACCAATCCGCTCCGCTGCGAAGGACTTG	6360
Db	6301	GAGCTCCGGAACCTATCCAGCAAGCGCGTTAAACCAATCCGCTCCGCTGCGAAGGACTTG	6360
QY	6361	CTGGAAGACTGAGACACCAATTGACACCAATCATGTCGCAAAATGAGGTTTCTGCG	6420
Db	6361	CTGGAAGACTGAGACACCAATTGACACCAATCATGTCGCAAAATGAGGTTTCTGCG	6420
QY	6421	GTCCAAACAGAGAGGGGGCGCGAGCTGCGCTTATCGTATTCAGATTTGGGG	6480
Db	6421	GTCCAAACAGAGAGGGGGCGCGAGCTGCGCTTATCGTATTCAGATTTGGGG	6480
QY	6481	GTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCGCTG	6540
Db	6481	GTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCGCTG	6540
QY	6541	ATGGGCTCTTCAAGGATTCGAATCTCTCTGGAAGCGGGTTCGAGTTTCTGGTGAAT	6600
Db	6541	ATGGGCTCTTCAAGGATTCGAATCTCTCTGGAAGCGGGTTCGAGTTTCTGGTGAAT	6600
QY	6601	GCTCTGAAGCGAAGAAATGSCCTATGGGCTTCGATATGACACCCGCTGTTGACTCA	6660
Db	6601	GCTCTGAAGCGAAGAAATGSCCTATGGGCTTCGATATGACACCCGCTGTTGACTCA	6660
QY	6661	ACGGTCACTGAGAAATGACATCCGCTGTGAGAGTCAATCTACCAATGTTGTGACTTGGCC	6720
Db	6661	ACGGTCACTGAGAAATGACATCCGCTGTGAGAGTCAATCTACCAATGTTGTGACTTGGCC	6720
QY	6721	CCGGAAGCCAGACAGGCCAATAAGGTGCTCAGAGCGGCTTTACATCGGGGCGCCCTG	6780
Db	6721	CCGGAAGCCAGACAGGCCAATAAGGTGCTCAGAGCGGCTTTACATCGGGGCGCCCTG	6780
QY	6781	ACTAATCTTAAAGGCGAGAACTGCGGCTATCGGCTATCGGAGCGGCTTACATCGGGGCGCCCTG	6840
Db	6781	ACTAATCTTAAAGGCGAGAACTGCGGCTATCGGCTATCGGAGCGGCTTACATCGGGGCGCCCTG	6840
QY	6841	ACCAGCTCGGTAATACCTCTACATGTTACTTGAAGCGCGCTCGCGCTGTCCAGCTGGC	6900
Db	6841	ACCAGCTCGGTAATACCTCTACATGTTACTTGAAGCGCGCTCGCGCTGTCCAGCTGGC	6900
QY	6901	AGCTCCAGACTGCAAGATGCTCGTATCGGAGCGAAGCTTGTCTGTTATCTGTGAAAGC	6960
Db	6901	AGCTCCAGACTGCAAGATGCTCGTATCGGAGCGAAGCTTGTCTGTTATCTGTGAAAGC	6960
QY	6961	GGGGGACCCAGAGACGAGCGGCTACGGGCTTCCAGGAGCTATGACTAGATAC	7020
Db	6961	GGGGGACCCAGAGACGAGCGGCTACGGGCTTCCAGGAGCTATGACTAGATAC	7020
QY	7021	TCTGCCCGCTGGGACCGCCCAACAGATACGACTTGGAGTTGATAATCATATGC	7080
Db	7021	TCTGCCCGCTGGGACCGCCCAACAGATACGACTTGGAGTTGATAATCATATGC	7080
QY	7081	TCTTCAATGTGTAGTGGGACGATGATCTGGCAAAAGGGTGTACTATCTCACCCGT	7140
Db	7081	TCTTCAATGTGTAGTGGGACGATGATCTGGCAAAAGGGTGTACTATCTCACCCGT	7140

QY	7141	GACCCCAACCCCCCTTTCGCGGGCTGCGTGGAGACAGCTAGACACACTCCAGTCAAT	7200
Db	7141	GACCCCAACCCCCCTTTCGCGGGCTGCGTGGAGACAGCTAGACACACTCCAGTCAAT	7200
QY	7201	TCTGGCTAGGCAACATCATGTATGCGCCACCTTGTGGCAAGATGATCCTGATG	7260
Db	7201	TCTGGCTAGGCAACATCATGTATGCGCCACCTTGTGGCAAGATGATCCTGATG	7260
QY	7261	ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACTTCAAAAAGCCCTAGATTGTGAG	7320
Db	7261	ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACTTCAAAAAGCCCTAGATTGTGAG	7320
QY	7321	ATCTAGGGGCTTGTACTCCATTTAGGACCACTTGAACCTCAGATCATTTCAACGACTC	7380
Db	7321	ATCTAGGGGCTTGTACTCCATTTAGGACCACTTGAACCTCAGATCATTTCAACGACTC	7380
QY	7381	CATGGCTTAGCGCATTTTTCACCTCCATAGTACTCTCCAGGTGAGATCAATAGGTTGGCT	7440
Db	7381	CATGGCTTAGCGCATTTTTCACCTCCATAGTACTCTCCAGGTGAGATCAATAGGTTGGCT	7440
QY	7441	TCATGCTCAGGAACTTGGGTTACCGCTTCCGAGTCTGGAGACATCGGGCCAGAACT	7500
Db	7441	TCATGCTCAGGAACTTGGGTTACCGCTTCCGAGTCTGGAGACATCGGGCCAGAACT	7500
QY	7501	GTCCGCTTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
Db	7501	GTCCGCTTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
QY	7561	AACCTGGGAGTAGGACCAAGCTCAAACTCACCTCCATCCCGCTGCTCCAGTTGGAT	7620
Db	7561	AACCTGGGAGTAGGACCAAGCTCAAACTCACCTCCATCCCGCTGCTCCAGTTGGAT	7620
QY	7621	TTATCCAGCTGTTCTGTTGCTGTACAGGGGGAGACATATATCACAGCTGTCTCGT	7680
Db	7621	TTATCCAGCTGTTCTGTTGCTGTACAGGGGGAGACATATATCACAGCTGTCTCGT	7680
QY	7681	GCCGACCCGCTGTTCTGTTGCTGTACAGGGGGAGACATATATCACAGCTGTCTCGT	7740
Db	7681	GCCGACCCGCTGTTCTGTTGCTGTACAGGGGGAGACATATATCACAGCTGTCTCGT	7740
QY	7741	CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT	7800
Db	7741	CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT	7800
QY	7801	TTTCCCTT	7860
Db	7801	TTTCCCTT	7860
QY	7861	TTTTTCTCTTT	7920
Db	7861	TTTTTCTCTTT	7920
QY	7921	TAGCTGTGAAGGCTCGTCCAGCGCTTGTACTCCAGAGTGTCTGATCTGGGCTCTGCG	7980
Db	7921	TAGCTGTGAAGGCTCGTCCAGCGCTTGTACTCCAGAGTGTCTGATCTGGGCTCTGCG	7980
QY	7981	AGATCAAGT 7989	
Db	7981	AGATCAAGT 7989	

## RESULT 2

ADP86264

ID ADP86264 standard; DNA; 11313 Bp.

XX AC ADP86264;

XX AC ADP86264;

DT 23-SEP-2004 (first entry)

DE Hepatitis C virus Con-1 replicon I377/NS3-3' plasmid DNA.

XX Hepatitis C virus; HCV; anti-HCV agent; HCV infection; therapy; plasmid; ds.



Db 1621 TGGCTCTCCTCAAGGTATTCAACAAGGGCTGAAGGATGCCAGAGGTACCCCATGTG 1680  
QY 1681 ATGGGATCTGATCTGGGGCTCGGTGACATGCTTTACATGTGTTAGTCGAGGTAAAA 1740  
Db 1681 ATGGGATCTGATCTGGGGCTCGGTGACATGCTTTACATGTGTTAGTCGAGGTAAAA 1740  
QY 1741 AACGTCCTAGGCCCCCGGAAACACAGGGGACGTGGTTTTCTTTGAAACACAGATAATACC 1800  
Db 1741 AACGTCCTAGGCCCCCGGAAACACAGGGGACGTGGTTTTCTTTGAAACACAGATAATACC 1800  
QY 1801 ATGGCGGCTATTAGGGCTACTCCAAACAGACGGAGGCTACTTTGGCTGCATCATCACT 1860  
Db 1801 ATGGCGGCTATTAGGGCTACTCCAAACAGACGGAGGCTACTTTGGCTGCATCATCACT 1860  
QY 1861 AGCCTCACGGCCGGGACAGGAACAGGTCCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
Db 1861 AGCCTCACGGCCGGGACAGGAACAGGTCCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
QY 1921 ACACAAATCTTTCTGGCGACTCGCTCAATGGCTGTGTGGACTGTCTATCATGTGGCC 1980  
Db 1921 ACACAAATCTTTCTGGCGACTCGCTCAATGGCTGTGTGGACTGTCTATCATGTGGCC 1980  
QY 1981 GGCTCAAGACCTTTGCCGCCCAAGAGGCCCAATCACCCAAATGTACACCAATGTGGAC 2040  
Db 1981 GGCTCAAGACCTTTGCCGCCCAAGAGGCCCAATCACCCAAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACTCTCGGCTGGCAAGCGCCCGCGGGCGGCTTCCTTGAACCAATGCACTGTC 2100  
Db 2041 CAGGACTCTCGGCTGGCAAGCGCCCGCGGGCGGCTTCCTTGAACCAATGCACTGTC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTTGGTCACGAGCATGCCGATGTCAATCCGGTGGCCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTTGGTCACGAGCATGCCGATGTCAATCCGGTGGCCGCGG 2160  
QY 2161 GGCAGCAGAGGGGAGCCTACTCTCCCGAGGCCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
Db 2161 GGCAGCAGAGGGGAGCCTACTCTCCCGAGGCCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
QY 2221 GGCAGTCTCGCTCTGCCCTCGGGCAGCTGTGGGCACTTTTCGGGCTGCGGTGTC 2280  
Db 2221 GGCAGTCTCGCTCTGCCCTCGGGCAGCTGTGGGCACTTTTCGGGCTGCGGTGTC 2280  
QY 2281 ACCCGAGGGGTTCCGAAGGCGGTGGACTTTGTATCCCGTCCAGCTATGGAAACCACTATG 2340  
Db 2281 ACCCGAGGGGTTCCGAAGGCGGTGGACTTTGTATCCCGTCCAGCTATGGAAACCACTATG 2340  
QY 2341 CGGTCCCGGCTTTCCAGGACAACTCGTCCCTCCGGCCGTACCGCAGACATTTCCAGGTG 2400  
Db 2341 CGGTCCCGGCTTTCCAGGACAACTCGTCCCTCCGGCCGTACCGCAGACATTTCCAGGTG 2400  
QY 2401 GCCCATCTACGCCCTTACTGTAGCGGCAAGCACTAAGTGCCTGCGGTGCGGTATGCA 2460  
Db 2401 GCCCATCTACGCCCTTACTGTAGCGGCAAGCACTAAGTGCCTGCGGTGCGGTATGCA 2460  
QY 2461 GCCCAAGGTATAAGGTGCTTGTCTTGAACCCCGTCCGTCCGCCCACTTAGGTTTCGGG 2520  
Db 2461 GCCCAAGGTATAAGGTGCTTGTCTTGAACCCCGTCCGTCCGCCCACTTAGGTTTCGGG 2520  
QY 2521 GGGTATATGTCTAAGGCAATGGTATCGACCCCTAACATCAGAACCGGGTAGGACCAATC 2580  
Db 2521 GGGTATATGTCTAAGGCAATGGTATCGACCCCTAACATCAGAACCGGGTAGGACCAATC 2580  
QY 2581 ACCAGGGTGCCCCATACGTAAGTCTCCACTATGGCAAGTTTTCGCCGACGGTGTTC 2640  
Db 2581 ACCAGGGTGCCCCATACGTAAGTCTCCACTATGGCAAGTTTTCGCCGACGGTGTTC 2640  
QY 2641 TCTGGGGGCGCTATGACATCAATAATGTGATGAGTGCACCTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGGCGCTATGACATCAATAATGTGATGAGTGCACCTCAACTGACTCGACCACT 2700  
QY 2701 ATCTGGGATCGGCACTCTTGGACCAAGCGGAGACGGCTGGAGCGCACTCTGCTGTG 2760  
Db 2701 ATCTGGGATCGGCACTCTTGGACCAAGCGGAGACGGCTGGAGCGCACTCTGCTGTG 2760

QY 2761 CTGCCACCGCTACGCTCCGGATCGGTCAACCGTCCACATCCAAACATCAGAGAGTG 2820  
Db 2761 CTGCCACCGCTACGCTCCGGATCGGTCAACCGTCCACATCCAAACATCAGAGAGTG 2820  
QY 2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGGCAAGCAATCCCAATCCAGACCAATC 2880  
Db 2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGGCAAGCAATCCCAATCCAGACCAATC 2880  
QY 2881 AAGGGGGGAGGACCTCATTTCTGTCATTTCCAAAGAAAGAAATGTAGCTCGCGCG 2940  
Db 2881 AAGGGGGGAGGACCTCATTTCTGTCATTTCCAAAGAAAGAAATGTAGCTCGCGCG 2940  
QY 2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTTGATGATCGGTC 3000  
Db 2941 AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTTGATGATCGGTC 3000  
QY 3001 ATACCACTAGCGGAGAGCTCATTTGTGTAGCAACCGACGCTCTAATGACGGGCTTTACC 3060  
Db 3001 ATACCACTAGCGGAGAGCTCATTTGTGTAGCAACCGACGCTCTAATGACGGGCTTTACC 3060  
QY 3061 GGGCATTTCCAGCTCAGTGTATCGACTGCAATACATGTGTACCCAGACAGTCTCAGC 3120  
Db 3061 GGGCATTTCCAGCTCAGTGTATCGACTGCAATACATGTGTACCCAGACAGTCTCAGC 3120  
QY 3121 CTGGACCCGACCTTCACTTGTAGACGACGACCGGTGCCAAGACGGGCTTCACTGCTG 3180  
Db 3121 CTGGACCCGACCTTCACTTGTAGACGACGACCGGTGCCAAGACGGGCTTCACTGCTG 3180  
QY 3181 CAGGGGGAGGACGAGCTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
Db 3181 CAGGGGGAGGACGAGCTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGGCCCTCGGGCATGTTTCGATTCCTCGTTCTGTGCGAGTCTATGACGGGGTGT 3300  
Db 3241 GAACGGCCCTCGGGCATGTTTCGATTCCTCGTTCTGTGCGAGTCTATGACGGGGTGT 3300  
QY 3301 GCTTGTAGAGCTACGCCCGCGGAGACCTCAGTTAGGTTTGGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTAGAGCTACGCCCGCGGAGACCTCAGTTAGGTTTGGGGCTTACCTAAACACA 3360  
QY 3361 CAGGGTTGGCCGTCTGCCAGACCAATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTC 3420  
Db 3361 CAGGGTTGGCCGTCTGCCAGACCAATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCTC 3420  
QY 3421 ACCCATAGAGCCCATTTCTTGTCCAGACTTAAGCAGCAGGAGACACTTCCCTTAC 3480  
Db 3421 ACCCATAGAGCCCATTTCTTGTCCAGACTTAAGCAGCAGGAGACACTTCCCTTAC 3480  
QY 3481 CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
Db 3481 CTGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
QY 3541 CAAATGTGGAAGTGTCTCATAGGCTAAAGCTTACGCTGCAAGGGCAAGCCCTGTGTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATAGGCTAAAGCTTACGCTGCAAGGGCAAGCCCTGTGTG 3600  
QY 3601 TATAGGCTGGGAGCGTTTCAAAACAGAGTTACTTACACACACCCCATACCAATAATC 3660  
Db 3601 TATAGGCTGGGAGCGTTTCAAAACAGAGTTACTTACACACACCCCATACCAATAATC 3660  
QY 3661 ATGGCATGATCTCGGCTGACCTGGAGTCTGTACGAGCACTGGGTGTGTAGGCGGA 3720  
Db 3661 ATGGCATGATCTCGGCTGACCTGGAGTCTGTACGAGCACTGGGTGTGTAGGCGGA 3720  
QY 3721 GTCTTAGAGCTCTGGCGCGTATTTGCTGTGACAAAGCAGCGTGGTCAATTTGGGCGAG 3780  
Db 3721 GTCTTAGAGCTCTGGCGCGTATTTGCTGTGACAAAGCAGCGTGGTCAATTTGGGCGAG 3780  
QY 3781 ATCACTTTGTCCGGAAGCCGCCCATCTTTCCGACAGGGAAGTCTTTTACCAGGAGTTC 3840  
Db 3781 ATCACTTTGTCCGGAAGCCGCCCATCTTTCCGACAGGGAAGTCTTTTACCAGGAGTTC 3840





Db 6001 ACAGGCGCCCTGATCAGCCCATGCGCTGCGGAGAAACCAAGCTGCCCATCATGACACTG 6060  
QY 6061 AGCAACTCTTTGCTCGGTCAACCAAACTTGGTCTATGCTACAAATCTCTCCGACGCGAAGC 6120  
Db 6061 AGCAACTCTTTGCTCGGTCAACCAAACTTGGTCTATGCTACAAATCTCTCCGACGCGAAGC 6120  
QY 6121 CTGCGGACAGAGAGTCACTTTGACAGACTGCAGTCTCTGACGACCACTACCGGGAC 6180  
Db 6121 CTGCGGACAGAGAGTCACTTTGACAGACTGCAGTCTCTGACGACCACTACCGGGAC 6180  
QY 6181 GTGCTCAAGAGATGAAGCGAAGGCTCCACAGTTTAAAGCTTAAAGCTTCTATCCGTTGGAG 6240  
Db 6181 GTGCTCAAGAGATGAAGCGAAGGCTCCACAGTTTAAAGCTTAAAGCTTCTATCCGTTGGAG 6240  
QY 6241 GAAGCCTGTAAGTGACGCCCCCACAATTCGGCCAGATCTAAATTTGGCTATGCGGCAAAAG 6300  
Db 6241 GAAGCCTGTAAGTGACGCCCCCACAATTCGGCCAGATCTAAATTTGGCTATGCGGCAAAAG 6300  
QY 6301 GACGTCGGAACTATCCAGCAGGCGCTTAAACCATCGCTCCGTTGGTGGAGGACTTG 6360  
Db 6301 GACGTCGGAACTATCCAGCAGGCGCTTAAACCATCGCTCCGTTGGTGGAGGACTTG 6360  
QY 6361 CTGGAAGACACTGAGACACCAATTGACACCAACCATCATGGCAAAAATGAGTTTCTGC 6420  
Db 6361 CTGGAAGACACTGAGACACCAATTGACACCAACCATCATGGCAAAAATGAGTTTCTGC 6420  
QY 6421 GTCCAAACAGAGAGGGGGCGCAAGCCAGTTCGCTTATCGTATTCACAGATTTGGGG 6480  
Db 6421 GTCCAAACAGAGAGGGGGCGCAAGCCAGTTCGCTTATCGTATTCACAGATTTGGGG 6480  
QY 6481 GTTCGTGTGTGAGAAATGCCCCTTTACGATGTGCTCCACCTCCCTCAGGCCGTG 6540  
Db 6481 GTTCGTGTGTGAGAAATGCCCCTTTACGATGTGCTCCACCTCCCTCAGGCCGTG 6540  
QY 6541 ATGGGCTCTTCATACGGAATCCAAATCTCTCTGGACAGCGGTTCGAGTTCTCTGGTGAAT 6600  
Db 6541 ATGGGCTCTTCATACGGAATCCAAATCTCTCTGGACAGCGGTTCGAGTTCTCTGGTGAAT 6600  
QY 6601 GCCTGAAAGCAGAGAAATGCCCTATGCGCTTCGCATATGACACCCGCTGTTTGTACTCA 6660  
Db 6601 GCCTGAAAGCAGAGAAATGCCCTATGCGCTTCGCATATGACACCCGCTGTTTGTACTCA 6660  
QY 6661 ACGGTCACTGAGATGACATCCGTGTGAGGAGTCAATCTACCAATGTTGACTGGCC 6720  
Db 6661 ACGGTCACTGAGATGACATCCGTGTGAGGAGTCAATCTACCAATGTTGACTGGCC 6720  
QY 6721 CCGGAAGCCAGACAGGCCATAGTTCGCTCACAGAGCGGCTTACATCGGGGCCCCCTG 6780  
Db 6721 CCGGAAGCCAGACAGGCCATAGTTCGCTCACAGAGCGGCTTACATCGGGGCCCCCTG 6780  
QY 6781 ACTAATCTAAAGGCGAGAACTGCGGCTATGCGCGTTCGCGCGAGCGGTGACTGACG 6840  
Db 6781 ACTAATCTAAAGGCGAGAACTGCGGCTATGCGCGTTCGCGCGAGCGGTGACTGACG 6840  
QY 6841 ACCAGCTCGGTAAATACCCTCACATGTTACTTGAAGCGCGCTGCGGCTGTCAGGCTGG 6900  
Db 6841 ACCAGCTCGGTAAATACCCTCACATGTTACTTGAAGCGCGCTGCGGCTGTCAGGCTGG 6900  
QY 6901 AAGCTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960  
Db 6901 AAGCTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960  
QY 6961 GCGGGACCCAGAGAGACGAGGCGGCTTACCGGCTTCACGAGGCTATGACTAGATAC 7020  
Db 6961 GCGGGACCCAGAGAGACGAGGCGGCTTACCGGCTTCACGAGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTCGGGACCGGCCCAACCAAGATACGACTTGGAGTTGATAATCATGCG 7080  
Db 7021 TCTGCCCCCTCGGGACCGGCCCAACCAAGATACGACTTGGAGTTGATAATCATGCG 7080  
QY 7081 TCTCCAAATGTGTCAGTTCGGCAGATGATCTGCGCAAAAGGCTGACTATCTCACCCTG 7140  
Db 7081 TCTCCAAATGTGTCAGTTCGGCAGATGATCTGCGCAAAAGGCTGACTATCTCACCCTG 7140

QY 7141 GACCCCAACACCCCCCTTGGCGGGCTGCGTGGAGACAGCTAGACACACTCCAGTCAAT 7200  
Db 7141 GACCCCAACACCCCCCTTGGCGGGCTGCGTGGAGACAGCTAGACACACTCCAGTCAAT 7200  
QY 7201 TCCTGCTAGGCAACATCATCATGATATGCCACCTTGTGGCAAGGATGATCCCTGATG 7260  
Db 7201 TCCTGCTAGGCAACATCATCATGATATGCCACCTTGTGGCAAGGATGATCCCTGATG 7260  
QY 7261 ACTCATTTTCTTCCATCCTTCTAGCTCAGGAACAACCTTGA AAAAGCCCTAGATTGCTAG 7320  
Db 7261 ACTCATTTTCTTCCATCCTTCTAGCTCAGGAACAACCTTGA AAAAGCCCTAGATTGCTAG 7320  
QY 7321 ATCTAGGGGCTGTTTATCTCCATTGAGCCACTTGACCTACCTCAGATCATTAACGACTC 7380  
Db 7321 ATCTAGGGGCTGTTTATCTCCATTGAGCCACTTGACCTACCTCAGATCATTAACGACTC 7380  
QY 7381 CATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
Db 7381 CATGCGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
QY 7441 TCATGCTCAGGAACCTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCCAGAAAT 7500  
Db 7441 TCATGCTCAGGAACCTTGGGGTACCGCCCTTTCGAGTCTGGAGACATCGGGCCAGAAAT 7500  
QY 7501 GTCCGCGTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Db 7501 GTCCGCGTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
QY 7561 AACTGGGAGTAAAGACCAAGCTCAAACTCACTCCAAATCCCGCTCGCTCCAGTTGGAT 7620  
Db 7561 AACTGGGAGTAAAGACCAAGCTCAAACTCACTCCAAATCCCGCTCGCTCCAGTTGGAT 7620  
QY 7621 TTATCCAGCTGTTTCTGTTGCTGTTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7680  
Db 7621 TTATCCAGCTGTTTCTGTTGCTGTTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7680  
QY 7681 GCCGACCCCGCTGTTTCTGTTGCTGTTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7740  
Db 7681 GCCGACCCCGCTGTTTCTGTTGCTGTTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7740  
QY 7741 CTACTCCCAACCGAGTAAAGCGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Db 7741 CTACTCCCAACCGAGTAAAGCGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
QY 7801 TTTCCCTTT 7860  
Db 7801 TTTCCCTTT 7860  
QY 7861 TTTTTCCTTT 7920  
Db 7861 TTTTTCCTTT 7920  
QY 7921 TAGCTGTGAAAGGTCCGTCAGCGCTTGTACTCAGAGAGTGTCTGACTACTGCTCTCTGC 7980  
Db 7921 TAGCTGTGAAAGGTCCGTCAGCGCTTGTACTCAGAGAGTGTCTGACTACTGCTCTCTGC 7980  
QY 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989

## RESULT 3

AAA98968  
ID AAA98968 standard; DNA; 7989 BP.  
XX  
AC AAA98968;  
XX  
DT 08-FEB-2001 (first entry)  
XX  
DE Hepatitis C virus DNA fragment SEQ ID NO: 4.  
XX  
KW Cell culture; therapy; infection; vaccine; diagnosis; gene therapy; ds.

xx Hepatitis C virus.  
 xx DE19915178-A1.  
 xx 05-OCT-2000.  
 xx 03-APR-1999; 99DE-01015178.  
 xx 03-APR-1999; 99DE-01015178.  
 xx (UYMA-) UNIV MAINZ GUTENBERG JOHANNES.  
 xx Bartenschlager R;  
 xx WPI; 2000-629140/61.  
 xx Cell culture system for hepatitis C virus, useful e.g. in screening for  
 xx therapeutic agents, comprises human hepatoma cells containing a viral RNA  
 xx construct that includes a selectable gene.  
 xx Claim 8; Page 37-43; 58pp; German.  
 xx This invention describes a novel Hepatitis C virus (HCV) cell culture  
 xx system comprising human hepatoma cells that contain an integrated HCV-RNA  
 xx construct (I). (i) contains the HCV-specific RNA segments 5'-NTR (non-  
 xx translated region), NS (non-structural), NS4A, NS4B, NS5A, NS5B and 3'-  
 xx NTR, and a selectable (marker) gene (II). The cell cultures, and/or (i),  
 xx are used to prepare, evaluate and/or test therapeutic and/or diagnostic  
 xx agents for HCV infections, and to prepare vaccines against HCV infection  
 xx (particularly preparation of attenuated HCV). The can also be used for  
 xx preparation of a liver-specific delivery system for gene therapy, and to  
 xx identify cells permissive for HCV replication. Virus RNA replicates  
 xx autonomously and with high efficiency in this cellular system, so that  
 xx variations in replication rates can be measured (for screening antiviral  
 xx agents) quantitatively or qualitatively, using standard laboratory  
 xx equipment. Efficient replication of HCV RNA is only achieved when the  
 xx specified RNA segments are present and when the transfected cells are  
 xx maintained under permanent selection pressure  
 xx Sequence 7989 BP; 1647 A; 2368 C; 2243 G; 1731 T; 0 U; 0 Other;  
 xx Query Match 100.0%; Score 7987.4; DB 3; Length 7989;  
 xx Best Local Similarity 100.0%; Pred. No. 0;  
 xx Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GCCAGCCCCCATGGGGGGGACACTCCACATAGATCACCCTCGTGGAGAACTACTG 60  
 DB 1 GCCAGCCCCCATGGGGGGGACACTCCACATAGATCACCCTCGTGGAGAACTACTG 60  
 QY 61 TCCTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGTGCGAGCTCCAGGAC 120  
 DB 61 TCCTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGTGCGAGCTCCAGGAC 120  
 QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180  
 DB 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180  
 QY 181 GACACCGGTGCTTCTTGATCAACCGCTCAATGCTGAGATTTGGCGTGCCTCC 240  
 DB 181 GACACCGGTGCTTCTTGATCAACCGCTCAATGCTGAGATTTGGCGTGCCTCC 240  
 QY 241 GCGAGATGCTAGCCGAGTAGTGTGGTCCGGAAGGCCCTTGTGTAATCCCTGATAGG 300  
 DB 241 GCGAGATGCTAGCCGAGTAGTGTGGTCCGGAAGGCCCTTGTGTAATCCCTGATAGG 300  
 QY 301 GTGCTTGGAGTGCCTCCGGAGGTCTCGTAGACCGTGCACCATGAGACGAACTCTAAAC 360  
 DB 301 GTGCTTGGAGTGCCTCCGGAGGTCTCGTAGACCGTGCACCATGAGACGAACTCTAAAC 360  
 QY 361 CTCAGAGAAACCAAAAGGCGCCCATGATTGACAGATGAGTTCGCCAGGTTCTC 420  
 DB 361 CTCAGAGAAACCAAAAGGCGCCCATGATTGACAGATGAGTTCGCCAGGTTCTC 420

QY 421 CGGCGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCACAAACAGACAAATCGGCTGCT 480  
 DB 421 CGGCGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCACAAACAGACAAATCGGCTGCT 480  
 QY 481 CTGATGCCCGCGTGTTCGGCTGTGAGCGAGGGGCGCCCGGTTCTTTTGTCAAGACCG 540  
 DB 481 CTGATGCCCGCGTGTTCGGCTGTGAGCGAGGGGCGCCCGGTTCTTTTGTCAAGACCG 540  
 QY 541 ACCTGTCCGGTGCCTGAAATGAACTGAGAGCGAGGCGAGCGGCTATCGTGGTGGGCA 600  
 DB 541 ACCTGTCCGGTGCCTGAAATGAACTGAGAGCGAGGCGAGCGGCTATCGTGGTGGGCA 600  
 QY 601 CGACGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGCG 660  
 DB 601 CGACGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGCG 660  
 QY 661 TGCTATTGGGCGAAGTGCAGGAGGATCTCCTGTCTATCTCACCTTCTCTGCGGAGA 720  
 DB 661 TGCTATTGGGCGAAGTGCAGGAGGATCTCCTGTCTATCTCACCTTCTCTGCGGAGA 720  
 QY 721 AAGTATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTGATCCGCTACTGCG 780  
 DB 721 AAGTATCCATCATGCTGATGCAATGCGCGGCTGCATACGCTTGATCCGCTACTGCG 780  
 QY 781 CATTTCGACCCACCAAGCGAATCATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
 DB 781 CATTTCGACCCACCAAGCGAATCATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
 QY 841 TTGTTCGATCAGATGATCTGGAAGAGCATAGGGGCTCGCGCCAGCCGAACTGTTCG 900  
 DB 841 TTGTTCGATCAGATGATCTGGAAGAGCATAGGGGCTCGCGCCAGCCGAACTGTTCG 900  
 QY 901 CCAGGCTCAAGGCGCGATGCGCGAGGAGATCTGCTGTAACCATCGGATGCT 960  
 DB 901 CCAGGCTCAAGGCGCGATGCGCGAGGAGATCTGCTGTAACCATCGGATGCT 960  
 QY 961 GCTTTCGCAATCATGCTGGAATGCGCGCTTTCTTGGATTCATGACGTGTCGCGG 1020  
 DB 961 GCTTTCGCAATCATGCTGGAATGCGCGCTTTCTTGGATTCATGACGTGTCGCGG 1020  
 QY 1021 TGGTGTGGGCGACCGCTTATCAGGACATAGCGTTGGCTACCGCTGATATTCCTGAAGAGC 1080  
 DB 1021 TGGTGTGGGCGACCGCTTATCAGGACATAGCGTTGGCTACCGCTGATATTCCTGAAGAGC 1080  
 QY 1081 TTGGCGGGAATGCGCTGACCGCTTCTGCTGCTTTAGGATTCGCGCTCCGATTCGC 1140  
 DB 1081 TTGGCGGGAATGCGCTGACCGCTTCTGCTGCTTTAGGATTCGCGCTCCGATTCGC 1140  
 QY 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTTCTTCTGAGTTTAAACAGACACACAG 1200  
 DB 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTTCTTCTGAGTTTAAACAGACACACAG 1200  
 QY 1201 GTTTCCTCTAGCGGATCAATCCGCGCTCTCCCTCCCGCCCTTAACGTTACTGCG 1260  
 DB 1201 GTTTCCTCTAGCGGATCAATCCGCGCTCTCCCTCCCGCCCTTAACGTTACTGCG 1260  
 QY 1261 CGAGCGCTTGGAAATGAGCGGCTGCGTTTCTATATGTTATTTTCCACATATTG 1320  
 DB 1261 CGAGCGCTTGGAAATGAGCGGCTGCGTTTCTATATGTTATTTTCCACATATTG 1320  
 QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTCGCGCTGTCTTCTTGGAGCATTCCT 1380  
 DB 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTCGCGCTGTCTTCTTGGAGCATTCCT 1380  
 QY 1381 AGGGGTCTTCCCTCTCGCCAAAGGAATGAAAGTCTGTTGAATGTCTGAGGAGCA 1440  
 DB 1381 AGGGGTCTTCCCTCTCGCCAAAGGAATGAAAGTCTGTTGAATGTCTGAGGAGCA 1440  
 QY 1441 GTTCCTCTGAGAGCTTCTTGAAGCAACACACGCTCTAGGACCTTTGAGGAGCGG 1500  
 DB 1441 GTTCCTCTGAGAGCTTCTTGAAGCAACACACGCTCTAGGACCTTTGAGGAGCGG 1500

1501 AACCCCCACCTGGCGCAGGTGCTCTGGGCCAAAGCAGCGTGTATAGATACACCT 1560  
1501 AACCCCCACCTGGCGCAGGTGCTCTGGGCCAAAGCAGCGTGTATAGATACACCT 1560  
1561 GCAAGGCGGCACAAACCCAGTGCAGCTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAA 1620  
1561 GCAAGGCGGCACAAACCCAGTGCAGCTTGTGAGTTGGATAGTTGTGGAAAGAGTCAAA 1620  
1621 TGGCTCTCCTCAAGCTATTCAACAAGGGGCTGAAGGATGCCAAGAGGTACCCCATGT 1680  
1621 TGGCTCTCCTCAAGCTATTCAACAAGGGGCTGAAGGATGCCAAGAGGTACCCCATGT 1680  
1681 ATGGGATCTCATCTGGGCTCTGGTGCACATGCTTTACATGTGTTTGTAGTCAAGTTAAAA 1740  
1681 ATGGGATCTCATCTGGGCTCTGGTGCACATGCTTTACATGTGTTTGTAGTCAAGTTAAAA 1740  
1741 AAGCTCTAGGCCGCCGGAACACAGGGGACGTGGTTTCCCTTTGAAAAACACGATATACC 1800  
1741 AAGCTCTAGGCCGCCGGAACACAGGGGACGTGGTTTCCCTTTGAAAAACACGATATACC 1800  
1801 ATGGGCTTATTAGGCTTACTCCCAACAGCGGAGGCTACTTGGCTGCATCATCACT 1860  
1801 ATGGGCTTATTAGGCTTACTCCCAACAGCGGAGGCTACTTGGCTGCATCATCACT 1860  
1861 AGCTCAACAGCGGACAGAAACAGTGCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
1861 AGCTCAACAGCGGACAGAAACAGTGCAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
1921 ACACAAATCTTCTGGGACCTCGTCAATGGCGTGTGTGGACTGTCTATCATGTGTC 1980  
1921 ACACAAATCTTCTGGGACCTCGTCAATGGCGTGTGTGGACTGTCTATCATGTGTC 1980  
1981 GGTCTAAGACCTTGGCGGCCCAAGGCCCCAATGACCCAAATGTACAAATGTGGAC 2040  
1981 GGTCTAAGACCTTGGCGGCCCAAGGCCCCAATGACCCAAATGTACAAATGTGGAC 2040  
2041 CAGGACCTGCTGGCTGGCAAGCGCCCGCGGCGCTTCTTGACACCATGCACTGC 2100  
2041 CAGGACCTGCTGGCTGGCAAGCGCCCGCGGCGCTTCTTGACACCATGCACTGC 2100  
2101 GGCAGCTCGGACCTTACTTGGTCAAGGACATGCGATGTCATTCGGGTGCGCGGCGG 2160  
2101 GGCAGCTCGGACCTTACTTGGTCAAGGACATGCGATGTCATTCGGGTGCGCGGCGG 2160  
2161 GCGCAGCAGGGGAGCCTACTCTCCCGAGCCCGTCTCTTACTTGAAGGCTCTTCG 2220  
2161 GCGCAGCAGGGGAGCCTACTCTCCCGAGCCCGTCTCTTACTTGAAGGCTCTTCG 2220  
2221 GCGGTCCTACTGCTCTGCCCTCGGGCAGCTGTGGGATCTTTCGGGCTGCGGTGTC 2280  
2221 GCGGTCCTACTGCTCTGCCCTCGGGCAGCTGTGGGATCTTTCGGGCTGCGGTGTC 2280  
2281 ACCGAGGGGTTGCAAGGCGGTGGACTTGTGTACCGCTCGAGTCTATGGAACCACTATG 2340  
2281 ACCGAGGGGTTGCAAGGCGGTGGACTTGTGTACCGCTCGAGTCTATGGAACCACTATG 2340  
2341 CGGTCCCGCTTTCACGACAACTCGTCCCTCGGCGGTACCGCAGACATTCAGGTG 2400  
2341 CGGTCCCGCTTTCACGACAACTCGTCCCTCGGCGGTACCGCAGACATTCAGGTG 2400  
2401 GCGCATCTACGCGCCCTACTGCTAGCGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
2401 GCGCATCTACGCGCCCTACTGCTAGCGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
2461 GCGCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGGCGCAACCTAGGTTTCGGG 2520  
2461 GCGCAAGGGTATAAGGTGCTTGTCTGAAACCGCTCGGCGCAACCTAGGTTTCGGG 2520  
2521 GGGTATGTCTAAGGCACATGGTATCGACCTTACATCAGAACCGGGGTAAAGCAATC 2580  
2521 GGGTATGTCTAAGGCACATGGTATCGACCTTACATCAGAACCGGGGTAAAGCAATC 2580  
2581 ACCAAGGTTGCCCCATCAGCTACTCCACCTATGCAAGTTTCTTGGCAGCGGTGTC 2640

2581 ACCAAGGTTGCCCCATCAGCTACTCCACCTATGCAAGTTTCTTGGCAGCGGTGTC 2640  
2641 TCTGGGGCGCCTATGACATCAATAATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
2641 TCTGGGGCGCCTATGACATCAATAATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
2701 ATCTGTGGCATCGGCACAGTCTCTGGACCAAGCGGAGACGGCTGGAGCGCACTCGTGTG 2760  
2701 ATCTGTGGCATCGGCACAGTCTCTGGACCAAGCGGAGACGGCTGGAGCGCACTCGTGTG 2760  
2761 CTGGCCACCGCTAGCGCTCGGGATCGGTTCACCGTGCACATCCAAATATCGAGAGGTG 2820  
2761 CTGGCCACCGCTAGCGCTCGGGATCGGTTCACCGTGCACATCCAAATATCGAGAGGTG 2820  
2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
2881 AAGGGGGGAGGCACTCTCAATTTCTGCCAATCCAAAGAAATGTGATGAGCTCGCGCG 2940  
2881 AAGGGGGGAGGCACTCTCAATTTCTGCCAATCCAAAGAAATGTGATGAGCTCGCGCG 2940  
2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATATCGCTC 3000  
2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATATCGCTC 3000  
3001 ATACCAACTAGCGGACAGTCTATTGTCTAGCAACGAGCGCTCTAATGACGGGCTTACC 3060  
3001 ATACCAACTAGCGGACAGTCTATTGTCTAGCAACGAGCGCTCTAATGACGGGCTTACC 3060  
3061 GCGATTTCCAGTCACTGATGCACTGCAATATGCTCACCAGACAGTCTGACTTCAGC 3120  
3061 GCGATTTCCAGTCACTGATGCACTGCAATATGCTCACCAGACAGTCTGACTTCAGC 3120  
3121 CTGGACCGACCTTACCAATGAGACGACGACCGTGCACAGACGCGGTGTACCGTGC 3180  
3121 CTGGACCGACCTTACCAATGAGACGACGACCGTGCACAGACGCGGTGTACCGTGC 3180  
3181 CAGCGCGAGGACGACTGCTAGGGGACGATGGGCATTTACAGGTTTCTGACTCCAGGA 3240  
3181 CAGCGCGAGGACGACTGCTAGGGGACGATGGGCATTTACAGGTTTCTGACTCCAGGA 3240  
3241 GAAAGCGCTCGGGCATGTTCCGATTTCTGCGAGTGTCTGCGAGTGTATACGCGGCTGT 3300  
3241 GAAAGCGCTCGGGCATGTTCCGATTTCTGCGAGTGTCTGCGAGTGTATACGCGGCTGT 3300  
3301 GCTTGTAGCAGCTCAGCGCCCGAGACCTCAGTTAGTTAGTTGCGGCTTACCTAAACACA 3360  
3301 GCTTGTAGCAGCTCAGCGCCCGAGACCTCAGTTAGTTAGTTGCGGCTTACCTAAACACA 3360  
3361 CCAGGTTGCGGCTGCGCAGGACCATCTGGAGTTCTGGGAGGCGTCTTTACAGGCTC 3420  
3361 CCAGGTTGCGGCTGCGCAGGACCATCTGGAGTTCTGGGAGGCGTCTTTACAGGCTC 3420  
3421 ACCCAATAGACGCCCATTTCTTCTGCCAGACTAAGCAGGACGAGGACAACTTCCCTTAC 3480  
3421 ACCCAATAGACGCCCATTTCTTCTGCCAGACTAAGCAGGACGAGGACAACTTCCCTTAC 3480  
3481 CTGTAGTACATACGAGCTACGTTGCGCAGGCTCAGGCTCAGCTCCTCATCTGTGGGAC 3540  
3481 CTGTAGTACATACGAGCTACGTTGCGCAGGCTCAGGCTCAGCTCCTCATCTGTGGGAC 3540  
3541 CAAATGTGGAAGTGTCTCATACGGCTAAGCCCTACGCTGACGGGCCAACGCCCTGTG 3600  
3541 CAAATGTGGAAGTGTCTCATACGGCTAAGCCCTACGCTGACGGGCCAACGCCCTGTG 3600  
3601 TATAGGCTGGGACCGGTTCAAAACGAGGTACTACACACACCCCATAAACCAATACATC 3660  
3601 TATAGGCTGGGACCGGTTCAAAACGAGGTACTACACACACCCCATAAACCAATACATC 3660  
3661 ATGGCATGCTGTGGCTGAGCTGAGGTGTCAGGACACTGGGTGCTGTAGGCGGA 3720

Db 3661 ATGSCATGTCGGCTGACCTGGAGTGTCTACGAGCACCTGGTGTGGTAGGGGA 3720  
Qy 3721 GTCCTAGAGCTCTGGCGCGGTATTTGCTGACAAAGCAGCGTGGTCAITTTGTGGGAGG 3780  
Db 3721 GTCCTAGAGCTCTGGCGCGGTATTTGCTGACAAAGCAGCGTGGTCAITTTGTGGGAGG 3780  
Qy 3781 ATCATCTTGTCCGGAAGCGCGCATCATTTCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGGAAGCGCGCATCATTTCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Qy 3841 GATGAGATGGAAGTGTGGCTTCACACCTCCCTTTACATCGAACAGGGAATCAGTCTGCC 3900  
Db 3841 GATGAGATGGAAGTGTGGCTTCACACCTCCCTTTACATCGAACAGGGAATCAGTCTGCC 3900  
Qy 3901 GAACAATTCAAACAGAGGGAATCGGTTGTGMAACAGGCCACCAAGCAAGCGGAGCT 3960  
Db 3901 GAACAATTCAAACAGAGGGAATCGGTTGTGMAACAGGCCACCAAGCAAGCGGAGCT 3960  
Qy 3961 GCTGCTCCGTTGGTGAATCCAAGTGGCGGACCTTCGAAGCTTCTGGCGGAAGCATATG 4020  
Db 3961 GCTGCTCCGTTGGTGAATCCAAGTGGCGGACCTTCGAAGCTTCTGGCGGAAGCATATG 4020  
Qy 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTTTCTCACTCTGCTTGGCAACCC 4080  
Db 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTTTCTCACTCTGCTTGGCAACCC 4080  
Qy 4081 GCGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCCGCTCACCAACCAACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCCGCTCACCAACCAACAT 4140  
Qy 4141 ACCCTCTGTTTAAATCTGGGGGATGGGTGGCGGCCCACTTGTCTCTCCACGCGT 4200  
Db 4141 ACCCTCTGTTTAAATCTGGGGGATGGGTGGCGGCCCACTTGTCTCTCCACGCGT 4200  
Qy 4201 GCTTCTGCTTTCGTAGGCGCGGATCGCTGAGCGGCTTTTGGAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGGCGCGGATCGCTGAGCGGCTTTTGGAGCATAGGCTTGGG 4260  
Qy 4261 AAGGTGCTTGGATATTTTGGCAGGTATGAGCAGGCGTGGAGGCGCTCTGTGGC 4320  
Db 4261 AAGGTGCTTGGATATTTTGGCAGGTATGAGCAGGCGTGGAGGCGCTCTGTGGC 4320  
Qy 4321 TTAAAGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAAACCTACTCTCCTGCT 4380  
Db 4321 TTAAAGTTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAAACCTACTCTCCTGCT 4380  
Qy 4381 ATCTCTCTCCCTGGCGCCTAGTCTGCGGGTCTGTGCGAGCGCATCTGCTGCGCAC 4440  
Db 4381 ATCTCTCTCCCTGGCGCCTAGTCTGCGGGTCTGTGCGAGCGCATCTGCTGCGCAC 4440  
Qy 4441 GTGGGCGCAGGGGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTCTGCGGG 4500  
Db 4441 GTGGGCGCAGGGGAGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTCTGCGGG 4500  
Qy 4501 GGTAAACAGCTCTCCCGCACGACATATGTGCTCTGAGAGGAGCGCTGACGACGTGTCACT 4560  
Db 4501 GGTAAACAGCTCTCCCGCACGACATATATGTGCTCTGAGAGGAGCGCTGACGACGTGTCACT 4560  
Qy 4561 CAGATCCTCTAGTCTTACCATCACTCAGCTGTGAAAGAGGTTTCCAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTAGTCTTACCATCACTCAGCTGTGAAAGAGGTTTCCAGTGGATCAAC 4620  
Qy 4621 GAGGACTGTCCAGCGCATGTCCGGCTCGTGGCTGAAGAGATGTTTGGGATGGATATGC 4680  
Db 4621 GAGGACTGTCCAGCGCATGTCCGGCTCGTGGCTGAAGAGATGTTTGGGATGGATATGC 4680  
Qy 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCTCTGCGCGGATTCGCGGA 4740  
Db 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCTCTGCGCGGATTCGCGGA 4740  
Qy 4741 GTCCCTCTCTTCTCATGTCAACGTGGGTACAGGAGTCTGGCGGGGCGACGCGATCATG 4800  
Db 4741 GTCCCTCTCTTCTCATGTCAACGTGGGTACAGGAGTCTGGCGGGGCGACGCGATCATG 4800

Qy 4801 CAAACACCTGCCCATGTGGAGCAGATCACCGACATGTGAAAAACGTTTCCATGAGG 4860  
Db 4801 CAAACACCTGCCCATGTGGAGCAGATCACCGACATGTGAAAAACGTTTCCATGAGG 4860  
Qy 4861 ATCGTGGGGCTAGGACCTGTAGTAAACGTFGGATGGAAATTTCCCATTAACGCGTAC 4920  
Db 4861 ATCGTGGGGCTAGGACCTGTAGTAAACGTFGGATGGAAATTTCCCATTAACGCGTAC 4920  
Qy 4921 ACCAGGGCCCTCGACGCCCTCCCGGCGCAAAATTTCTAGGGGCTGTGGCGGGTG 4980  
Db 4921 ACCAGGGCCCTCGACGCCCTCCCGGCGCAAAATTTCTAGGGGCTGTGGCGGGTG 4980  
Qy 4981 GCTGCTGAGGAGTACGTGGAGGTTACGCGGTGGGGGATTTTCCACTACGTGACGGGATG 5040  
Db 4981 GCTGCTGAGGAGTACGTGGAGGTTACGCGGTGGGGGATTTTCCACTACGTGACGGGATG 5040  
Qy 5041 ACCACTGACAAAGTAAAGTTCGAGTTCGAGGTTCCGGCCCGCAATTTCTTACAGAGTG 5100  
Db 5041 ACCACTGACAAAGTAAAGTTCGAGTTCGAGGTTCCGGCCCGCAATTTCTTACAGAGTG 5100  
Qy 5101 GATGGGTGCGGTTGCAACAGGTACGCTCCAGCTGCAAAACCCCTCTACGGGAGGAGTCT 5160  
Db 5101 GATGGGTGCGGTTGCAACAGGTACGCTCCAGCTGCAAAACCCCTCTCTACGGGAGGAGTCT 5160  
Qy 5161 ACATTTCTGTTGCGGCTCAATCAATACCTGTTGGGTCAACAGTCCCATGCGAGCCGAA 5220  
Db 5161 ACATTTCTGTTGCGGCTCAATCAATACCTGTTGGGTCAACAGTCCCATGCGAGCCGAA 5220  
Qy 5221 CCGAGCTAGCAGTGTCTCACTTCCATGCTACGAGTCCGAGTCCCATGCGGAGGAGG 5280  
Db 5221 CCGAGCTAGCAGTGTCTCACTTCCATGCTACGAGTCCGAGTCCCATGCGGAGGAGG 5280  
Qy 5281 GCTAAGCTTAGCTGCGGCGAGGAGTCTCCCGCTCTTGGCGAGTCACTACGTAGCCAG 5340  
Db 5281 GCTAAGCTTAGCTGCGGCGAGGAGTCTCCCGCTCTTGGCGAGTCACTACGTAGCCAG 5340  
Qy 5341 CTGCTGCGCTTCTTGAAGCAACATGCACTACCTCGTCACTCCCGCTCATGCTCCCGAGCGTAC 5400  
Db 5341 CTGCTGCGCTTCTTGAAGCAACATGCACTACCTCGTCACTCCCGCTCATGCTCCCGAGCGTAC 5400  
Qy 5401 CTCATCGAGGCAACCTCTGTGGCGGAGAGATGGCGGGGAAATCACCCGCTGGAG 5460  
Db 5401 CTCATCGAGGCAACCTCTGTGGCGGAGAGATGGCGGGGAAATCACCCGCTGGAG 5460  
Qy 5461 TCAGAAATAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAAGCGGAGGATGAG 5520  
Db 5461 TCAGAAATAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAAGCGGAGGATGAG 5520  
Qy 5521 AGGGAAGTATCCGTTCCGCGGAGATCTCTGCGAGGTCCAGGAAATTCCTCGAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCGCGGAGATCTCTGCGAGGTCCAGGAAATTCCTCGAGCGATG 5580  
Qy 5581 CCCATATGGGCAACGCGCGGATTAACACCTCCACTGTTAGAGTCTTGGAGGACCGGAC 5640  
Db 5581 CCCATATGGGCAACGCGCGGATTAACACCTCCACTGTTAGAGTCTTGGAGGACCGGAC 5640  
Qy 5641 TAGGTCCCTCCAGTGTACAGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA 5700  
Db 5641 TAGGTCCCTCCAGTGTACAGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA 5700  
Qy 5701 CTTCCAGGAGGAGGAGGAGGCTGTCTGTGTCAGAAATCTACCGTGTCTTCTGCTTGGGG 5760  
Db 5701 CTTCCAGGAGGAGGAGGAGGCTGTCTGTGTCAGAAATCTACCGTGTCTTCTGCTTGGGG 5760  
Qy 5761 GAGCTGCGCACAAAGACCTTCGCGAGCTCCGAATCTGTCGAGGATCCGAGCGGACCGCA 5820  
Db 5761 GAGCTGCGCACAAAGACCTTCGCGAGCTCCGAATCTGTCGAGGATCCGAGCGGACCGCA 5820  
Qy 5821 ACGGCTCTCTGACGAGCGCTCCCGAGCGGAGCGGAGTCCGAGCTTGAAGTGTGATC 5880  
Db 5821 ACGGCTCTCTGACGAGCGCTCCCGAGCGGAGCGGAGTCCGAGCTTGAAGTGTGATC 5880

Qy	5881	TCCTCCATGCCCCCTTGTAGGGGGAGCGGGGATCCGATCTCAGCAGCGGTCTTGG	5941
Db	5881	TCCTCCATGCCCCCTTGTAGGGGGAGCGGGGATCCGATCTCAGCAGCGGTCTTGG	5940
Qy	5941	TCACCCGTAGCGAGGAGGTAGTGAGGAGCTGCTCTGCTGCTCATGTCCTACACATGG	6000
Db	5941	TCACCCGTAGCGAGGAGGTAGTGAGGAGCTGCTCTGCTGCTCATGTCCTACACATGG	6000
Qy	6001	ACAGCGCCCTTGATACGCCATCGCTCGCGAGGAAAACCAAGCTGCCCATCAATGCACTG	6060
Db	6001	ACAGCGCCCTTGATACGCCATCGCTCGCGAGGAAAACCAAGCTGCCCATCAATGCACTG	6060
Qy	6061	AGAACTCTTTGTCCTCGTCAACAACTTGGTCTATGCTACAACTCTCGCAGCGAAGC	6120
Db	6061	AGAACTCTTTGTCCTCGTCAACAACTTGGTCTATGCTACAACTCTCGCAGCGAAGC	6120
Qy	6121	CTGGCGCAGAGAAGAGTCACTTTTGACAGACTGCAAGTCTTGGACGACCACTACCGGGAC	6180
Db	6121	CTGGCGCAGAGAAGAGTCACTTTTGACAGACTGCAAGTCTTGGACGACCACTACCGGGAC	6180
Qy	6181	GTGCTCAAGGAGATGAAGGGGAAAGGCGTCCACAGCTTAAGGCTAAACTTCTATCCGTGGAG	6240
Db	6181	GTGCTCAAGGAGATGAAGGGGAAAGGCGTCCACAGCTTAAGGCTAAACTTCTATCCGTGGAG	6240
Qy	6241	GAAGCCTGTGAAGTGAAGCGCCCAATTCGGCCAGATCTAAATTTGGCTATGGGCAAG	6300
Db	6241	GAAGCCTGTGAAGTGAAGCGCCCAATTCGGCCAGATCTAAATTTGGCTATGGGCAAG	6300
Qy	6301	GAGTCCGGAACTATCCAGCAGGCGGTAAACCACTCCGCTCCGTGTGGAGGACTTG	6360
Db	6301	GAGTCCGGAACTATCCAGCAGGCGGTAAACCACTCCGCTCCGTGTGGAGGACTTG	6360
Qy	6361	CTGGAAGACACTGAGACACCAATTTGACCAACCAATCATGSCAAAATAAGGTTTCTGCG	6420
Db	6361	CTGGAAGACACTGAGACACCAATTTGACCAACCAATCATGSCAAAATAAGGTTTCTGCG	6420
Qy	6421	GTCCAAACGAGAAAGGGGGCGGAAAGCCAGCTGCGCTTATCGTATTTCCAGATTTGGGG	6480
Db	6421	GTCCAAACGAGAAAGGGGGCGGAAAGCCAGCTGCGCTTATCGTATTTCCAGATTTGGGG	6480
Qy	6481	GTTCTGTGTGGAGAAATGGCCCTTTAGATGTGTCTCCACCTCCCTCAGGCGGTG	6540
Db	6481	GTTCTGTGTGGAGAAATGGCCCTTTAGATGTGTCTCCACCTCCCTCAGGCGGTG	6540
Qy	6541	ATGGGCTCTTATACGGATTCCAATATCTCTCTGGACAGCGGTTCGAGTTCCTGTGTGAAT	6600
Db	6541	ATGGGCTCTTATACGGATTCCAATATCTCTCTGGACAGCGGTTCGAGTTCCTGTGTGAAT	6600
Qy	6601	GCCTGGAAGCGAAGAAATGCCCTATGCGGCTTCGCATATGACACCGCGTGTTTGACTCA	6660
Db	6601	GCCTGGAAGCGAAGAAATGCCCTATGCGGCTTCGCATATGACACCGCGTGTTTGACTCA	6660
Qy	6661	ACGGTCACTCAGAAATGACATCCGTGTGAGGAGTCAATCTACCAATGTTGTGACTGGCC	6720
Db	6661	ACGGTCACTCAGAAATGACATCCGTGTGAGGAGTCAATCTACCAATGTTGTGACTGGCC	6720
Qy	6721	CCGGAAGCCAGACAGGCCATAAGGTCCCTCAGAGCGGTTCATCGGGGCGCCCTTG	6780
Db	6721	CCGGAAGCCAGACAGGCCATAAGGTCCCTCAGAGCGGTTCATCGGGGCGCCCTTG	6780
Qy	6781	ACTAAATTTAAAGGCGAAGAACTCGCGGTATCGCGCGTTCGAGCGGTGTGACTGACG	6840
Db	6781	ACTAAATTTAAAGGCGAAGAACTCGCGGTATCGCGCGTTCGAGCGGTGTGACTGACG	6840
Qy	6841	ACCAGCTGCGGTAAATACCTTCAATGTTACTTGAAGCGCGCTCGGCGCTGTGAGCTGG	6900
Db	6841	ACCAGCTGCGGTAAATACCTTCAATGTTACTTGAAGCGCGCTCGGCGCTGTGAGCTGG	6900
Qy	6901	AAGCTCCAGGACTGCACATGCTCGTATGCGGAGAGCACTTCTGCTTACTGTGAAGC	6960
Db	6901	AAGCTCCAGGACTGCACATGCTCGTATGCGGAGAGCACTTCTGCTTACTGTGAAGC	6960
Qy	6961	GCGGGGACCCAGAGGACGAGCGAGCCTACGGGCTTTCAGGAGGCTATGACTAGATAC	7020

6961	Db	CGCGGGACCCAAAGGACGAGCGAGCCTACGGGSCCTTACGGAGGCTATGACTAGATAC	702
7021	QY	TCTGCCCCCTTGGGAGCCGCCCAAACACAGATACGACTTGGAGTTGATAACATCATGC	7080
7021	Db	TCTGCCCCCTTGGGAGCCGCCCAAACACAGATACGACTTGGAGTTGATAACATCATGC	7080
7081	QY	TCCTCCAAATGTTCAGTCGCGCAGATGCATCTGGCAAAAGGGTGTAATCTCTCACCCGT	7140
7081	Db	TCCTCCAAATGTTCAGTCGCGCAGATGCATCTGGCAAAAGGGTGTAATCTCTCACCCGT	7140
7141	QY	GACCCCAACCCCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT	7200
7141	Db	GACCCCAACCCCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT	7200
7201	QY	TCCTGGCTAGGCAACATCATGTATGCGGCCACTTGTGGGCAAGGATGATCCTGATG	7260
7201	Db	TCCTGGCTAGGCAACATCATGTATGCGGCCACTTGTGGGCAAGGATGATCCTGATG	7260
7261	QY	ACTCATTTCTTCTCCATCCTTCTAGCTCAGGAACAACCTTCAAAAAGCCCTAGATTTGTCAG	7320
7261	Db	ACTCATTTCTTCTCCATCCTTCTAGCTCAGGAACAACCTTCAAAAAGCCCTAGATTTGTCAG	7320
7321	QY	ATCTACGGGCGCTGTACTCCATAGGCCACTTGAGACCTCCTCAGATCATTTCAACGATC	7380
7321	Db	ATCTACGGGCGCTGTACTCCATAGGCCACTTGAGACCTCCTCAGATCATTTCAACGATC	7380
7381	QY	CATGGCCCTTAGCCGATTTTTCACCTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT	7440
7381	Db	CATGGCCCTTAGCCGATTTTTCACCTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT	7440
7441	QY	TCATGCGCTCAGGAACTTGGGGTACCGCCCTTCGGAGTCTGGAGACATCGGGCCAGAACT	7500
7441	Db	TCATGCGCTCAGGAACTTGGGGTACCGCCCTTCGGAGTCTGGAGACATCGGGCCAGAACT	7500
7501	QY	GTCGGCGCTAGGCTACGTGCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
7501	Db	GTCGGCGCTAGGCTACGTGCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
7561	QY	AACTGGCGCATGAGGACCAAGCTCAAACTCACCTCCAATCCGGCTCGGTCCCAAGTTGGAT	7620
7561	Db	AACTGGCGCATGAGGACCAAGCTCAAACTCACCTCCAATCCGGCTCGGTCCCAAGTTGGAT	7620
7621	QY	TTATCCAGCTGGTTCTGTTGCTGTATCAGCGGGGAGACATATATCACAGCCTGTCTCGT	7680
7621	Db	TTATCCAGCTGGTTCTGTTGCTGTATCAGCGGGGAGACATATATCACAGCCTGTCTCGT	7680
7681	QY	GCCGACCCCGCTGGTTCATGTGGTGCTACTCTACTTCTGTAGGGTAGGCATCTAT	7740
7681	Db	GCCGACCCCGCTGGTTCATGTGGTGCTACTCTACTTCTGTAGGGTAGGCATCTAT	7740
7741	QY	CTACTCCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCCATCTCTGTTTT	7800
7741	Db	CTACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCCATCTCTGTTTT	7800
7801	QY	TTTCCCTTTTTTTTTTCTT	7860
7801	Db	TTTCCCTTTTTTTTTTCTT	7860
7861	QY	TTTTTCTCTTTTTTTTTTCTTTTTCTTTCTTTCTGTTGCTCCATCTTAGCCCTAGTCACGGC	7920
7861	Db	TTTTTCTCTTTTTTTTTTCTTTTTCTTTCTTTCTGTTGCTCCATCTTAGCCCTAGTCACGGC	7920
7921	QY	TAGCTGTGAAGGTCGCTGAGCGGCTTGACTGCAGAGAGTGCTGATCTCTGCTCTCTGCTC	7980
7921	Db	TAGCTGTGAAGGTCGCTGAGCGGCTTGACTGCAGAGAGTGCTGATCTCTGCTCTCTGCTC	7980
7981	QY	AGATCAAGT	7989
7981	Db	AGATCAAGT	7989

AAD25326  
ID AAD25326 standard; cDNA; 7989 BP.  
XX AC AAD25326;  
XX 12-MAR-2002 (first entry)  
XX DE Hepatitis C virus (HCV) adaptive replicon V mutant cDNA.  
XX Hepatitis C virus; HCV; transfection; infection; virus neutralisation;  
KW gene therapy; vaccine; immunoprotection; hepatotropic; virucide; liver;  
KW adaptive replicon V; mutant; ss.  
XX Hepatitis C virus.  
OS Synthetic.  
XX  
FH Location/Qualifiers  
FT CDS 4642..5982  
FT /\*tag= a  
FT /product= "NSSA variant of HCV adaptive replicon v"  
FT /note= "CDS does not include both start and stop codon"  
FT /partial  
FT mutation replace(5320, G)  
FT /\*tag= b  
XX  
PN WO200189364-A2.  
XX 29-NOV-2001.  
XX  
XX 23-MAY-2001; 2001WO-US016822.  
XX  
XX 23-MAY-2000; 2000US-00576989.  
XX (UNIW ) UNIV WASHINGTON.  
XX  
XX Rice CM, Blight KJ;  
XX  
XX WPI; 2002-066755/09.  
XX P-PSDB; AAE15724.  
XX  
XX Hepatitis C virus variants having greater transfection efficiency and  
PT ability to survive subpassage, useful as a vaccine for immunizing primate  
PT to the virus, comprise non-naturally occurring viral sequences.  
XX  
XX Disclosure; Page 80-83; 174pp; English.  
XX  
XX The invention relates to Hepatitis C virus (HCV) variants which include  
CC polynucleotides comprising non-naturally occurring HCV sequence and HCV  
CC variants that have a transfection efficiency and ability to survive  
CC subpassage greater than HCV that have wild-type polypeptide coding  
CC regions. The polynucleotides of the invention are useful for identifying  
CC a cell line that is permissive for infection with HCV and detecting  
CC replication of HCV in cells of the cell line. They are also useful for  
CC testing a compound for anti-viral properties and for inhibiting HCV  
CC infection. They are also useful for the generation of defined HCV virus  
CC stocks to develop in vitro and in vivo assays for virus neutralisation,  
CC attachment, penetration and entry, structure/function studies on HCV  
CC proteins and RNA elements and identification of new antiviral targets, a  
CC systematic survey of cell culture systems and conditions to identify  
CC those that support wild-type and variant HCV RNA replication and particle  
CC release, production of adaptive HCV variants capable of more efficiency  
CC replication in cell culture, production of HCV variants with altered  
CC tissue or species tropism, establishment of alternative animal models for  
CC inhibitor evaluation including those supporting HCV variant replication,  
CC development of cell-free HCV replication assays, production of  
CC immunogenic HCV particles for vaccination, engineering of attenuated HCV  
CC derivatives as possible vaccine candidates, engineering of attenuated or  
CC defective HCV derivatives for expression of heterologous gene products  
CC for gene therapy and vaccine applications and for utilisation of the HCV  
CC glycoproteins for targeted delivery of therapeutic agents to the liver  
CC or other cell types with appropriate receptors. Vaccine comprising these  
CC sequences is useful for inducing immunoprotection to HCV in a primate.  
XX The present sequence is Hepatitis C virus (HCV) adaptive replicon V

CC mutant cDNA. This sequence is generated by the mutation g to t at  
CC position 5320 of HCVreplbBartMan/Avail cDNA  
XX  
XX Sequence 7989 BP; 1647 A; 2369 C; 2241 G; 1732 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 7987.4; DB 6; Length 7989;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCCAGCCCCCGAATTGGGGGCGACATTCACCATAGATCACTCCCTGTGAGGAATCTGTG 60  
Db 1 GCCAGCCCCCGAATTGGGGGCGACATTCACCATAGATCACTCCCTGTGAGGAATCTGTG 60  
QY 61 TCTTCAGCGCAGAAAGCGTCTAGCCATGCGCTTAGTATGAGTGTCTGTCAGCTCCAGGAC 120  
Db 61 TCTTCAGCGCAGAAAGCGTCTAGCCATGCGCTTAGTATGAGTGTCTGTCAGCTCCAGGAC 120  
QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATGCCAG 180  
Db 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATGCCAG 180  
QY 181 GACGACCGGTCCTTCTTTGGATCAACCCGCTCAATGCTCGAGATTGGGCGTCCGCC 240  
Db 181 GACGACCGGTCCTTCTTTGGATCAACCCGCTCAATGCTCGAGATTGGGCGTCCGCC 240  
QY 241 GCGAGACTGCTAGCGGAGTGTGGTCTGCGGAAGGCTTGTGGTACTGCTCTGATAGG 300  
Db 241 GCGAGACTGCTAGCGGAGTGTGGTCTGCGGAAGGCTTGTGGTACTGCTCTGATAGG 300  
QY 301 GTGCTTGGAGTGGCCCGGAGGTTCTGTAGCCGTCACCATGAGCAGCAATCTCTAAAC 360  
Db 301 GTGCTTGGAGTGGCCCGGAGGTTCTGTAGCCGTCACCATGAGCAGCAATCTCTAAAC 360  
QY 361 CTCGAAGAAAACCAAGGCGCGCATGATGACCAAGATGATGACGAGGTTGTCACGAGTTCTC 420  
Db 361 CTCGAAGAAAACCAAGGCGCGCATGATGACCAAGATGATGACGAGGTTGTCACGAGTTCTC 420  
QY 421 CGGCGCTTGGTGGAGAGGCTATTGGCTATGACTGGGCGACAAACAGACAATCGGCTGT 480  
Db 421 CGGCGCTTGGTGGAGAGGCTATTGGCTATGACTGGGCGACAAACAGACAATCGGCTGT 480  
QY 481 CTGATGCGCGGTTGTCGGTGTGAGCGAGGGCGCGGTTCTTTTGTCAAGACCG 540  
Db 481 CTGATGCGCGGTTGTCGGTGTGAGCGAGGGCGCGGTTCTTTTGTCAAGACCG 540  
QY 541 ACCTGTCGGTGCCTGTAATGAACTGCAAGGAGCGAGCGGCGGTATCGTGGCTGGCCA 600  
Db 541 ACCTGTCGGTGCCTGTAATGAACTGCAAGGAGCGAGCGGCGGTATCGTGGCTGGCCA 600  
QY 601 CGACGGGCTTCTTGGCGAGCTGTGCTCGACGTTGTCTCACTGAAGCGGAAGGAGTGGC 660  
Db 601 CGACGGGCTTCTTGGCGAGCTGTGCTCGACGTTGTCTCACTGAAGCGGAAGGAGTGGC 660  
QY 661 TGCTATTGGGGAAGTGGCGGGGAGGATCTCTGTCTCTCACCTTGTCTCTGCGGAGA 720  
Db 661 TGCTATTGGGGAAGTGGCGGGGAGGATCTCTGTCTCTCACCTTGTCTCTGCGGAGA 720  
QY 721 AAGTATCCATCATGCTGATGCAATGCGGCGCTGATACGCTTGTATCGGCTACTGCG 780  
Db 721 AAGTATCCATCATGCTGATGCAATGCGGCGCTGATACGCTTGTATCGGCTACTGCG 780  
QY 781 CATTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
Db 781 CATTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAGCGGTC 840  
QY 841 TTGTCGATCAGATGATCTGACGAGAGCATCAGGGGCTCGGCGAGCGCACTGTTTCG 900  
Db 841 TTGTCGATCAGATGATCTGACGAGAGCATCAGGGGCTCGGCGAGCGCACTGTTTCG 900  
QY 901 CCAGGCTCAAGCGCGCATGCGCGGAGGATCTCGTCTGTGACCCATGCGGATGCGCT 960  
Db 901 CCAGGCTCAAGCGCGCATGCGCGGAGGATCTCGTCTGTGACCCATGCGGATGCGCT 960



Qy 961 GCTTGGCCGAATATCATGTTGAAATAGCCGCTTTCTGGATTCATCGACTGTGCGCGC 1020  
Db 961 GCTTGGCCGAATATCATGTTGAAATAGCCGCTTTCTGGATTCATCGACTGTGCGCGC 1020  
Qy 1021 TGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTCCTGAAGAGC 1080  
Db 1021 TGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGTGATATTCCTGAAGAGC 1080  
Qy 1081 TTGGCGGGAATGGCTGACCGCTTCCTCGTCTTACGGTATGCGCGCTCCGATTCGC 1140  
Db 1081 TTGGCGGGAATGGCTGACCGCTTCCTCGTCTTACGGTATGCGCGCTCCGATTCGC 1140  
Qy 1141 AGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCCTCTGAGTTTAAACAGACACACAGC 1200  
Db 1141 AGCGCATCGCCTTCTATCGCCTTCTTGACGAGTTCCTCTGAGTTTAAACAGACACACAGC 1200  
Qy 1201 GTTTCCTCTAGCGGATCAATTCGGCCCTCTCCCTCCGCCCTTAAAGTTACTGGC 1260  
Db 1201 GTTTCCTCTAGCGGATCAATTCGGCCCTCTCCCTCCGCCCTTAAAGTTACTGGC 1260  
Qy 1261 CGAAGCGCTTGAATAGCCGCTGTGCTTCTATATGTTATTTTCACCATATG 1320  
Db 1261 CGAAGCGCTTGAATAGCCGCTGTGCTTCTATATGTTATTTTCACCATATG 1320  
Qy 1321 CCGTCTTTTGGCAATGTGAGGCGCCGGAACCTGGCCCTGTCTTCTTGACGAGCATTCCT 1380  
Db 1321 CCGTCTTTTGGCAATGTGAGGCGCCGGAACCTGGCCCTGTCTTCTTGACGAGCATTCCT 1380  
Qy 1381 AGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAATGTCGTGAAGGAAGCA 1440  
Db 1381 AGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAATGTCGTGAAGGAAGCA 1440  
Qy 1441 GTTCTCTGGAAGTCTTCTGAAGCAACCAACGTCTGTAGGACCTTTGAGGACGCGG 1500  
Db 1441 GTTCTCTGGAAGTCTTCTGAAGCAACCAACGTCTGTAGGACCTTTGAGGACGCGG 1500  
Qy 1501 AACCCTCCACTGTGGCAAGGTGCTCTGGGCCAAAGCAACGTGTATAGATACACCT 1560  
Db 1501 AACCCTCCACTGTGGCAAGGTGCTCTGGGCCAAAGCAACGTGTATAGATACACCT 1560  
Qy 1561 GAAAGCGGCAAAACCCAGTGCACGTGTGAGTTGATGTTGTGAAAGAGTCAAA 1620  
Db 1561 GAAAGCGGCAAAACCCAGTGCACGTGTGAGTTGATGTTGTGAAAGAGTCAAA 1620  
Qy 1621 TGGCTCTCAAGCTATTCACAGGCGCTGAAGATCCCAAGAGTACCCATTGT 1680  
Db 1621 TGGCTCTCAAGCTATTCACAGGCGCTGAAGATCCCAAGAGTACCCATTGT 1680  
Qy 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGGTTAAA 1740  
Db 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGGTTAAA 1740  
Qy 1741 AAGCTTAGCCCGCCGAAACAGGGACGTGTTTCTTTGAAACACGATAATACC 1800  
Db 1741 AAGCTTAGCCCGCCGAAACAGGGACGTGTTTCTTTGAAACACGATAATACC 1800  
Qy 1801 ATGGCGCTATTACGCGCTACTCCACAGACCGGAGGCTACTTGGTGATCATCACT 1860  
Db 1801 ATGGCGCTATTACGCGCTACTCCACAGACCGGAGGCTACTTGGTGATCATCACT 1860  
Qy 1861 AGCCTCACAGCGGGACAGGAACAGGTGAGGGGAGGTCCAGTGGTCTCCACCGCA 1920  
Db 1861 AGCCTCACAGCGGGACAGGAACAGGTGAGGGGAGGTCCAGTGGTCTCCACCGCA 1920  
Qy 1921 ACACAATCTTCTGCGGACCTCGTCAATGGCGTGTGTTGAGCTGTCTATCATGTGCC 1980  
Db 1921 ACACAATCTTCTGCGGACCTCGTCAATGGCGTGTGTTGAGCTGTCTATCATGTGCC 1980  
Qy 1981 GGTCAAGACCCCTTCCCGCCCAAGAGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
Db 1981 GGTCAAGACCCCTTCCCGCCCAAGAGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
Qy 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGCGGCTTCTTTGACACCATGCACCTGC 2100

Db 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGGGCGGCTTCTTTGACACCATGCACCTGC 2100  
Qy 2101 GGCAGCTCGGACCTTTACTTGGTCAAGGATGATCCGATGTCATTCGGTGGCGGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTGGTCAAGGATGATCCGATGTCATTCGGTGGCGGCGG 2160  
Qy 2161 GGCACACAGGAGGAGCTTACTCTCCGCCAGGCCGCTCTCTACTTGAAGGGCTCTTCG 2220  
Db 2161 GGCACACAGGAGGAGCTTACTCTCCGCCAGGCCGCTCTCTACTTGAAGGGCTCTTCG 2220  
Qy 2221 GGCAGCTCGGACCTTCTGCGCTCGGSCACGCTGTGGGCTCTTTGGGCTGCGGTGTC 2280  
Db 2221 GGCAGCTCGGACCTTCTGCGCTCGGSCACGCTGTGGGCTCTTTGGGCTGCGGTGTC 2280  
Qy 2281 ACCGAGGGGTGCAAGGAGGCTGACTTTGTACCGTTCGAGTCTATGGAAACCACTATG 2340  
Db 2281 ACCGAGGGGTGCAAGGAGGCTGACTTTGTACCGTTCGAGTCTATGGAAACCACTATG 2340  
Qy 2341 CGGTCCCGGTCTTACCGGACAACTCGTCCCTCCGCCGTACCGCAGACATTCAGGTG 2400  
Db 2341 CGGTCCCGGTCTTACCGGACAACTCGTCCCTCCGCCGTACCGCAGACATTCAGGTG 2400  
Qy 2401 GGCATCTACAGCCCTTACTGCTAGCGGCAAGGACATTAAGTTCGCGCTGCGTATGCA 2460  
Db 2401 GGCATCTACAGCCCTTACTGCTAGCGGCAAGGACATTAAGTTCGCGCTGCGTATGCA 2460  
Qy 2461 GCCAAGGGTATAGGTGCTTGTCTGAAACCGGCTCGCGCCACCCCTAGGTTTCGGG 2520  
Db 2461 GCCAAGGGTATAGGTGCTTGTCTGAAACCGGCTCGCGCCACCCCTAGGTTTCGGG 2520  
Qy 2521 GCGTATATGCTAAGGCACATGCTATCGCCCTAAACATCAGAACCGGGTAAAGACCATC 2580  
Db 2521 GCGTATATGCTAAGGCACATGCTATCGCCCTAAACATCAGAACCGGGTAAAGACCATC 2580  
Qy 2581 ACCAGGGTGGCCCGCATCAGTACTCACCTATGGCAAGTTCTTTCGCGAGGTTGCTG 2640  
Db 2581 ACCAGGGTGGCCCGCATCAGTACTCACCTATGGCAAGTTCTTTCGCGAGGTTGCTG 2640  
Qy 2641 TCTGGGGCGCTTATGACATCAATATGATGAGTGCCACTCAACTGACTCGACCAT 2700  
Db 2641 TCTGGGGCGCTTATGACATCAATATGATGAGTGCCACTCAACTGACTCGACCAT 2700  
Qy 2701 ATCTGGGCATCGGCACAGTCTTGGACCAAGCGGAGACGCGCTGGAGCGGCTCGTGTG 2760  
Db 2701 ATCTGGGCATCGGCACAGTCTTGGACCAAGCGGAGACGCGCTGGAGCGGCTCGTGTG 2760  
Qy 2761 CTGCGCACCGCTACGCTCGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG 2820  
Db 2761 CTGCGCACCGCTACGCTCGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG 2820  
Qy 2821 GCTCTCTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Db 2821 GCTCTCTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Qy 2881 AAGGGGGGAGGACCTCATTTCTGCCATTCAGAGAAATGTGATGAGTTCGCGCG 2940  
Db 2881 AAGGGGGGAGGACCTCATTTCTGCCATTCAGAGAAATGTGATGAGTTCGCGCG 2940  
Qy 2941 AAGCTGTCCGCGCTCGGACTCAATGTGTAGCATATTACCGGGCGCTTGTATCCGTC 3000  
Db 2941 AAGCTGTCCGCGCTCGGACTCAATGTGTAGCATATTACCGGGCGCTTGTATCCGTC 3000  
Qy 3001 ATACCAACTAGCGGAGACGTCAATGTGTAGCAACGAGCGCTCTAATGACGGCTTTACC 3060  
Db 3001 ATACCAACTAGCGGAGACGTCAATGTGTAGCAACGAGCGCTCTAATGACGGCTTTACC 3060  
Qy 3061 GCGATTTTCGACTCAGTGTGACTGCAATACATGTGTCAACAGACAGTTCGACTTCAG 3120  
Db 3061 GCGATTTTCGACTCAGTGTGACTGCAATACATGTGTCAACAGACAGTTCGACTTCAG 3120  
Qy 3121 CTGGACCGGACCTTCAACATTCAGACGAGCAGCGGTGCCAAGACGCGGTGTACGCTCG 3180

Db 3121 CTGGACCGACCTTACCACTTGGACGACGACCGGTGCCACAGACGCGGTGTCAAGCTCG 3180  
Qy 3181 CAGCGCCAGGACAGCACTGGTAGGGGACAGGATGGGCATTTACAGTTTGTGACTCCAGGA 3240  
Db 3181 CAGCGCGAGGACAGCACTGGTAGGGGACAGGATGGGCATTTACAGTTTGTGACTCCAGGA 3240  
Qy 3241 GAAAGCCCTCGGGCATCTCGATTCTTCGGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
Db 3241 GAAAGCCCTCGGGCATCTCGATTCTTCGGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
Qy 3301 GCTTGGTACGAGCTCACGCCGCCCGAGACCTCAGTTAGTGTTCGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGGTACGAGCTCACGCCGCCCGAGACCTCAGTTAGTGTTCGGGCTTACCTAAACACA 3360  
Qy 3361 CCAGGGTTGCCGTCTGCCAGAACCATCTGAGATTTCTGGGAGAGGTCTTTACAGGCTC 3420  
Db 3361 CCAGGGTTGCCGTCTGCCAGAACCATCTGAGATTTCTGGGAGAGGTCTTTACAGGCTC 3420  
Qy 3421 ACCCACATAGAGCCCATTTCTGTCCAGACTAAGCAGGACAGGACAACTTCCCTAC 3480  
Db 3421 ACCCACATAGAGCCCATTTCTGTCCAGACTAAGCAGGACAGGACAACTTCCCTAC 3480  
Qy 3481 CTGGTAGCATACAGGCTACGGTGTGGCGCAGGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
Db 3481 CTGGTAGCATACAGGCTACGGTGTGGCGCAGGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
Qy 3541 CAAATGTGAAGTGTCTCATAGGCTAAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAGCCT 3600  
Db 3541 CAAATGTGAAGTGTCTCATAGGCTAAAGCCTAAGCCTAAGCCTAAGCCTAAGCCTAAGCCT 3600  
Qy 3601 TATAGCTGGAGCGCTTCAAAACAGAGTTACTTACCACACACCCCAATCAAAATACATC 3660  
Db 3601 TATAGCTGGAGCGCTTCAAAACAGAGTTACTTACCACACACCCCAATCAAAATACATC 3660  
Qy 3661 ATGGCATGATGTCGGCTGACCTGGAGTGTGTGACGAGACCTGGTGTGTGGCGGA 3720  
Db 3661 ATGGCATGATGTCGGCTGACCTGGAGTGTGTGACGAGACCTGGTGTGTGGCGGA 3720  
Qy 3721 GTCTAGAGCTTGGCGGGTATTCGCTGAACAGGAGAGTGTGTGTGTGGCGAG 3780  
Db 3721 GTCTAGAGCTTGGCGGGTATTCGCTGAACAGGAGAGTGTGTGTGTGGCGAG 3780  
Qy 3781 ATCACTCTGCGGAAGCGGCCATCATTTCCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Db 3781 ATCACTCTGCGGAAGCGGCCATCATTTCCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Qy 3841 GATGATGGAAGAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCAAGCTCGCC 3900  
Db 3841 GATGATGGAAGAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCAAGCTCGCC 3900  
Qy 3901 GAACAAATTCAGAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAGGAGGCT 3960  
Db 3901 GAACAAATTCAGAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAGGAGGCT 3960  
Qy 3961 GCTGCTCCGTGGTGAATTCAGAGTGCAGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGTGGTGAATTCAGAGTGCAGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Qy 4021 TGGAAATTTATCAGCGGGATACAAATTTTAGAGGCTTGTCCACTCTGCTGCGCAACCCC 4080  
Db 4021 TGGAAATTTATCAGCGGGATACAAATTTTAGAGGCTTGTCCACTCTGCTGCGCAACCCC 4080  
Qy 4081 GCGATAGCATCATGATGCAATTCAGAGCTCTTACACAGCGGCTCACACCCACAT 4140  
Db 4081 GCGATAGCATCATGATGCAATTCAGAGCTCTTACACAGCGGCTCACACCCACAT 4140  
Qy 4141 ACCCTCTGTTTAACTCTCGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGGCT 4200  
Db 4141 ACCCTCTGTTTAACTCTCGGGGATGGTGGCGGCCCAACTTGTCTCTCCAGGCT 4200  
Qy 4201 GCTTCTGCTTTCGTAGCGCGCGCATGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGCGCGCGCATGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260

Qy 4261 AAGTGTCTTGGATATTTTGGCAGGTTATGGACAGGGTGGCAGGCGCGCTCGTGGCC 4320  
Db 4261 AAGTGTCTTGGATATTTTGGCAGGTTATGGACAGGGTGGCAGGCGCGCTCGTGGCC 4320  
Qy 4321 TTTAAGTTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGT 4380  
Db 4321 TTTAAGTTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGT 4380  
Qy 4381 ATCCTCTCCCTGGCGCCCTAGTCTGCGGGTCTGTGCGGACGATGCTGCTGGCAC 4440  
Db 4381 ATCCTCTCCCTGGCGCCCTAGTCTGCGGGTCTGTGCGGACGATGCTGCTGGCAC 4440  
Qy 4441 GTGGGCCAGGGAGGGGCTGTGCACTGGATGAACCGGCTGATAGCGTTCCCTTCGCGG 4500  
Db 4441 GTGGGCCAGGGAGGGGCTGTGCACTGGATGAACCGGCTGATAGCGTTCCCTTCGCGG 4500  
Qy 4501 GGTAAACCACTCTCCCAACGCACTATGTGCTGAGAGGACGCTGCGAGCAGTGTCACT 4560  
Db 4501 GGTAAACCACTCTCCCAACGCACTATGTGCTGAGAGGACGCTGCGAGCAGTGTCACT 4560  
Qy 4561 CAGATCCTCTTAGTCTTACCACCTCAGCTGCTGAAGAGGCTTCCACAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTTAGTCTTACCACCTCAGCTGCTGAAGAGGCTTCCACAGTGGATCAAC 4620  
Qy 4621 GAGGACTGTCTCCACGCCATGTCTCCGGCTCGTGGCTAAGAGATGTTTGGATTTGATGC 4680  
Db 4621 GAGGACTGTCTCCACGCCATGTCTCCGGCTCGTGGCTAAGAGATGTTTGGATTTGATGC 4680  
Qy 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCTGCGCGGATTTCCGCGGA 4740  
Db 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCTGCGCGGATTTCCGCGGA 4740  
Qy 4741 GTCCCTCTCTCATGTCAACGTTGGGTACAGGGAGTCTGGCGGGGACGCGATCATG 4800  
Db 4741 GTCCCTCTCTCATGTCAACGTTGGGTACAGGGAGTCTGGCGGGGACGCGATCATG 4800  
Qy 4801 CAAACCACTGCCCCATGTGAGCACAGATCACCGACATGTGAAACACGGTTCCATGAGG 4860  
Db 4801 CAAACCACTGCCCCATGTGAGCACAGATCACCGACATGTGAAACACGGTTCCATGAGG 4860  
Qy 4861 ATCGTGGGCTTAGGACCTGTAGTAACAGTGGCTGGAACATTTCCCATTAACGCGTAC 4920  
Db 4861 ATCGTGGGCTTAGGACCTGTAGTAACAGTGGCTGGAACATTTCCCATTAACGCGTAC 4920  
Qy 4921 ACCAGGGCCCTGACAGCCCTCCCGCGCCAAATTTCTAGGCGCTGTGCGCGGTG 4980  
Db 4921 ACCAGGGCCCTGACAGCCCTCCCGCGCCAAATTTCTAGGCGCTGTGCGCGGTG 4980  
Qy 4981 GCTGCTGAGAGTACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGGGATG 5040  
Db 4981 GCTGCTGAGAGTACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGGGATG 5040  
Qy 5041 ACCACTGAACAACTAAAGTCCCGTGTGAGTTCCGGCCCCCGAATTTCTACAGAGTG 5100  
Db 5041 ACCACTGAACAACTAAAGTCCCGTGTGAGTTCCGGCCCCCGAATTTCTACAGAGTG 5100  
Qy 5101 GATGGGTGGGTGACAGGTACGCTCCAGCGTGAACCCCTCTACGGGAGGAGTTC 5160  
Db 5101 GATGGGTGGGTGACAGGTACGCTCCAGCGTGAACCCCTCTACGGGAGGAGTTC 5160  
Qy 5161 ACATTCCTGCTGGGTCAATCAATACCTGTTTGGGTCAAGCTCCATCGAGGCCGAA 5220  
Db 5161 ACATTCCTGCTGGGTCAATCAATACCTGTTTGGGTCAAGCTCCATCGAGGCCGAA 5220  
Qy 5221 CCGGAGCTAGAGTGTCTTCCATCTCACGCCCTCCACATTAACGCGGAGACG 5280  
Db 5221 CCGGAGCTAGAGTGTCTTCCATCTCACGCCCTCCACATTAACGCGGAGACG 5280  
Qy 5281 GCTAAGCTAGGCTGCGCAGGGATCTCCCTCTCTTGGCCAGCTCATCAGCTTAGCCAG 5340  
Db 5281 GCTAAGCTAGGCTGCGCAGGGATCTCCCTCTCTTGGCCAGCTCATCAGCTTAGCCAG 5340





Db 661 TGGTATTGGCGAAGTCGCGGGCAGGATCTCCCTGTCACTCACTTGTCTCTCGCGAGA 720  
Qy  
Db 721 AAGTATCGATCATGGCTGATGCAATCGCGGGCTGCAATAGCTTTGATCCGGCTACCTGCC 780  
721 AAGTATCGATCATGGCTGATGCAATCGCGGGCTGCAATAGCTTTGATCCGGCTACCTGCC 780  
Qy 781 CATTCGACCAACCAAGCGGAACATCGCATCGAGCGAGCACGTAATCGATGGAAGCGGTC 840  
Db 781 CATTCGACCAACCAAGCGGAACATCGCATCGAGCGAGCACGTAATCGATGGAAGCGGTC 840  
Qy 841 TTGTCGATCAGGATGATCTGGAAGAGACATCAGGGGCTCGCGCCAGCGCAACTGTTTCG 900  
Db 841 TTGTCGATCAGGATGATCTGGAAGAGACATCAGGGGCTCGCGCCAGCGCAACTGTTTCG 900  
Qy 901 CAGGCTCAAGCGCGGATGATCGGCGAGGAGATCTCGTGTGACCCATCGGCGATCCCT 960  
Db 901 CAGGCTCAAGCGCGGATGATCGGCGAGGAGATCTCGTGTGACCCATCGGCGATCCCT 960  
Qy 961 GCTTGGCGAATATCATGGTGAAGATCAGGGGCTCGCGCCAGCGCAACTGTTTCG 1020  
Db 961 GCTTGGCGAATATCATGGTGAAGATCAGGGGCTCGCGCCAGCGCAACTGTTTCG 1020  
Qy 1021 TGGGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGC 1080  
Db 1021 TGGGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGC 1080  
Qy 1081 TTGGCGGCAATGGGCTGACCGCTTCTCGTGTGTTACGGTATCGCGCTCCCGATTCGC 1140  
Db 1081 TTGGCGGCAATGGGCTGACCGCTTCTCGTGTGTTACGGTATCGCGCTCCCGATTCGC 1140  
Qy 1141 AGCGATCGCTTCTATCGCTTCTTGACGAGTCTCTGACGAGTCTCTGAGTTAAACAGACCAAGC 1200  
Db 1141 AGCGATCGCTTCTATCGCTTCTTGACGAGTCTCTGACGAGTCTCTGAGTTAAACAGACCAAGC 1200  
Qy 1201 GTTTCCTCTAGCGGGATCAATTCCGCGCTCTCCCTCCCGCCCTAAACGTTACTGGC 1260  
Db 1201 GTTTCCTCTAGCGGGATCAATTCCGCGCTCTCCCTCCCGCCCTAAACGTTACTGGC 1260  
Qy 1261 CGAAGCGCTTGAATAAGGCGGTGCGTTTGTATATGTTATTTTCCACCATATTG 1320  
Db 1261 CGAAGCGCTTGAATAAGGCGGTGCGTTTGTATATGTTATTTTCCACCATATTG 1320  
Qy 1321 CGCTCTTTTGGCAATGTAGGCGCCCGGAACCTGCGGCTGTCTCTTGACGACATTCCT 1380  
Db 1321 CGCTCTTTTGGCAATGTAGGCGCCCGGAACCTGCGGCTGTCTCTTGACGACATTCCT 1380  
Qy 1381 AGGGTCTTTTCCCTCTCGCCAAAGGAAATGCAAGGTCTGTTGAATGTGTTGAAGGAGCA 1440  
Db 1381 AGGGTCTTTTCCCTCTCGCCAAAGGAAATGCAAGGTCTGTTGAATGTGTTGAAGGAGCA 1440  
Qy 1441 GTTCCTCTGGAAGCTTCTTGAAGACAAAACAACTCTGTAGCGACCCCTTTGACGCGAGCGG 1500  
Db 1441 GTTCCTCTGGAAGCTTCTTGAAGACAAAACAACTCTGTAGCGACCCCTTTGACGCGAGCGG 1500  
Qy 1501 AACCCCGCCACCTCGGACAGTGCCTCTCGGCGCAAGCCAGCTGATAAGNATACCT 1560  
Db 1501 AACCCCGCCACCTCGGACAGTGCCTCTCGGCGCAAGCCAGCTGATAAGNATACCT 1560  
Qy 1561 GCAAGGCGGCAACCCCGAGTGCACCTGTGTGAGTTGATGTTGTTGAAGAGTCAAA 1620  
Db 1561 GCAAGGCGGCAACCCCGAGTGCACCTGTGTGAGTTGATGTTGTTGAAGAGTCAAA 1620  
Qy 1621 TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGGATGCCAGAGGTACCCCATTTGT 1680  
Db 1621 TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGGATGCCAGAGGTACCCCATTTGT 1680  
Qy 1681 ATGGGATCTGATCTGGGCTCGGTGCAATGCTTTTACATGTTTGTAGTCGAGGTTAAAA 1740  
Db 1681 ATGGGATCTGATCTGGGCTCGGTGCAATGCTTTTACATGTTTGTAGTCGAGGTTAAAA 1740  
Qy 1741 AACGCTTAGGCGCCCGGACCAACCGGAGCGTGGTTTCTTTTGAAGAAACAGGATATACC 1800  
Db

Db 1741 AACGCTTAGGCGCCCGGAAACCAACGCGGACGTGGTTCCTTTTGAAAAACAGATATACC 1800  
Qy 1801 ATGGCGCTATTAGCGCTTACTCCAAACAGACGCGAGCCCTACTTGGTGTGATCATCACT 1860  
Db 1801 ATGGCGCTATTAGCGCTTACTCCAAACAGACGCGAGCCCTACTTGGTGTGATCATCACT 1860  
Qy 1861 AGCTTACAGCGCGGAGACAGAAACAGGTGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
Db 1861 AGCTTACAGCGCGGAGACAGAAACAGGTGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
Qy 1921 ACACAATCTTTCTGGGACCTGGTCAATGGCGTGTGTTGGACTGTCTATCATGGTGCC 1980  
Db 1921 ACACAATCTTTCTGGGACCTGGTCAATGGCGTGTGTTGGACTGTCTATCATGGTGCC 1980  
Qy 1981 GGCTCAAGAACCTTGGCGGCCCCAAAGGGCCCCAATCAACCCAAATGTACACCAATGTGAGC 2040  
Db 1981 GGCTCAAGAACCTTGGCGGCCCCAAAGGGCCCCAATCAACCCAAATGTACACCAATGTGAGC 2040  
Qy 2041 CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGGCGGTTCTTGACACCAATGACCTGC 2100  
Db 2041 CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGGCGGTTCTTGACACCAATGACCTGC 2100  
Qy 2101 GGCAGCTCGGACCTTTTACITTTGGTCAAGGACATGCCGATGTCAATCCGGTGGCGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTTACITTTGGTCAAGGACATGCCGATGTCAATCCGGTGGCGCGCGG 2160  
Qy 2161 GGCAGCAGCAGGGGAGCCCTACTCTCCCCCAGGCGCCCTCTCTACTTTGAAAGGGCTCTTCG 2220  
Db 2161 GGCAGCAGCAGGGGAGCCCTACTCTCCCCCAGGCGCCCTCTCTACTTTGAAAGGGCTCTTCG 2220  
Qy 2221 GCGGCTCCACTGTCTGCGGCTCGGGGCGAGCTGTGGGCAATCTTCGGGCTGCGGTGTC 2280  
Db 2221 GCGGCTCCACTGTCTGCGGCTCGGGGCGAGCTGTGGGCAATCTTCGGGCTGCGGTGTC 2280  
Qy 2281 ACCCGAGGGGTTCCGAAGCGGTGGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCCGAGGGGTTCCGAAGCGGTGGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
Qy 2341 CGGTCCCGGCTTTCACGGACAACTCGTCCCTCCCGGCGGTACCGCAGACATTTCCAGGTG 2400  
Db 2341 CGGTCCCGGCTTTCACGGACAACTCGTCCCTCCCGGCGGTACCGCAGACATTTCCAGGTG 2400  
Qy 2401 GCCCATCTACGCCCTTCTGTTAGCGGCAAGAGCACTAAGGTGCGGCTGGTATGCA 2460  
Db 2401 GCCCATCTACGCCCTTCTGTTAGCGGCAAGAGCACTAAGGTGCGGCTGGTATGCA 2460  
Qy 2461 GCCCAAGGGTATAAGGTGCTTCTGCTGAACCCGTGCGTCCCGCCACCCCTAGGTTTCGGG 2520  
Db 2461 GCCCAAGGGTATAAGGTGCTTCTGCTGAACCCGTGCGTCCCGCCACCCCTAGGTTTCGGG 2520  
Qy 2521 GCGTATATCTTAAGGCACATGATCGACCCCTAAATCAGAAACCGGGGTAAAGGACATC 2580  
Db 2521 GCGTATATCTTAAGGCACATGATCGACCCCTAAATCAGAAACCGGGGTAAAGGACATC 2580  
Qy 2581 ACCAAGGGTCCCGCATACGTAATCTCACTATGGAAGGTTCCTTCGCGAGGTGGTTCG 2640  
Db 2581 ACCAAGGGTCCCGCATACGTAATCTCACTATGGAAGGTTCCTTCGCGAGGTGGTTCG 2640  
Qy 2641 TCTGGGGGCGCTATGACATCATATATGATGAGTGCCTCACTCACTGACCTGACCACT 2700  
Db 2641 TCTGGGGGCGCTATGACATCATATATGATGAGTGCCTCACTCACTGACCTGACCACT 2700  
Qy 2701 ATCTGGGCGCATCGGACAGTCTCTGGAACCAAGCGGAGAGCGGCTGGAGCGGCACTCGTCG 2760  
Db 2701 ATCTGGGCGCATCGGACAGTCTCTGGAACCAAGCGGAGAGCGGCTGGAGCGGCACTCGTCG 2760  
Qy 2761 CTCGCCACCGCTACCGCTCCGGGATCGGTTCACCGTGCACATCCAAACATCCAGGAGGTG 2820  
Db 2761 CTCGCCACCGCTACCGCTCCGGGATCGGTTCACCGTGCACATCCAAACATCCAGGAGGTG 2820  
Qy 2821 GCTCTGTCAGCACTGGAGAAATCCCTTTTATGCAAGCCATCCCATTCGAGACCATC 2880  
Db 2821 GCTCTGTCAGCACTGGAGAAATCCCTTTTATGCAAGCCATCCCATTCGAGACCATC 2880

QY	2881	AAGGGGGGAGGACCTCATTTTCTGCCATTCACAAAGAAATGTGATGAGCTGCGCGG	2940
Db	2881	AAGGGGGGAGGACCTCATTTTCTGCCATTCACAAAGAAATGTGATGAGCTGCGCGG	2940
QY	2941	AAGCTGTCGGGCTCGGACTCAATGCTGTAGCAVATTACCGGGGCTTGATGATCCGTC	3000
Db	2941	AAGCTGTCGGGCTCGGACTCAATGCTGTAGCAVATTACCGGGGCTTGATGATCCGTC	3000
QY	3001	ATACCAACTAGCGGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATGACGGGCTTTACC	3060
Db	3001	ATACCAACTAGCGGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATGACGGGCTTTACC	3060
QY	3061	GGGATTTTCAGTCTAGTGTGATCGATGCAATGCAATGATGTGATCCAGACAGTGTGACTCAGC	3120
Db	3061	GGGATTTTCAGTCTAGTGTGATCGATGCAATGCAATGATGTGATCCAGACAGTGTGACTCAGC	3120
QY	3121	CTGGACCCGACCTTCACCATTTAGAGACGACCGGTGCCAAGACGCGGTGTACGCTCG	3180
Db	3121	CTGGACCCGACCTTCACCATTTAGAGACGACCGGTGCCAAGACGCGGTGTACGCTCG	3180
QY	3181	CAGCGGCGAGGACGACTGTGTAGGGGACGATGGGCATTTACAGTTTGTGACTCCAGGA	3240
Db	3181	CAGCGGCGAGGACGACTGTGTAGGGGACGATGGGCATTTACAGTTTGTGACTCCAGGA	3240
QY	3241	GAACGGCCCTCGGGCATGTTGATTTCTCGGTCTGTGCGAGTGTATGACGCGGGCTGT	3300
Db	3241	GAACGGCCCTCGGGCATGTTGATTTCTCGGTCTGTGCGAGTGTATGACGCGGGCTGT	3300
QY	3301	GCTTGCTAGAGCTCACGCGCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA	3360
Db	3301	GCTTGCTAGAGCTCACGCGCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA	3360
QY	3361	CCAGGGTTGCCGCTCTGCGAGACCATCTGGAGTTCTGGGAGAGGCTCTTTACAGGCTTC	3420
Db	3361	CCAGGGTTGCCGCTCTGCGAGACCATCTGGAGTTCTGGGAGAGGCTCTTTACAGGCTTC	3420
QY	3421	ACCCATAGAGCCCATTTCTGTCAGATCTAGAGTCTAGGCTTACGAGGAGGAGCAACTTCCCTAC	3480
Db	3421	ACCCATAGAGCCCATTTCTGTCAGATCTAGAGTCTAGGCTTACGAGGAGGAGCAACTTCCCTAC	3480
QY	3481	CTGGTAGCATACAGGCTACGCTGTGCGCAGGGCTCAGGCTCCACCTCCATCTGGGAC	3540
Db	3481	CTGGTAGCATACAGGCTACGCTGTGCGCAGGGCTCAGGCTCCACCTCCATCTGGGAC	3540
QY	3541	CAAAATGTGGAAGTGTCTATACGGCTTAAAGCTTACGCTGCAAGGCGCAACGCGCTGTG	3600
Db	3541	CAAAATGTGGAAGTGTCTATACGGCTTAAAGCTTACGCTGCAAGGCGCAACGCGCTGTG	3600
QY	3601	TATAGCTGGGAGCCGTTCAAAACGAGGTACTACACACACCCCATACCAATACATC	3660
Db	3601	TATAGCTGGGAGCCGTTCAAAACGAGGTACTACACACACCCCATACCAATACATC	3660
QY	3661	ATGGCATGATGTGCGCTGACCTGAGGTCGTGACAGAGACCTGGGTGTGTTAGGCGGA	3720
Db	3661	ATGGCATGATGTGCGCTGACCTGAGGTCGTGACAGAGACCTGGGTGTGTTAGGCGGA	3720
QY	3721	GTCTTAGCAGCTGTGCGCGGTATTTGCTTGAACAGGAGCGTGTGATGTTGGGCGAG	3780
Db	3721	GTCTTAGCAGCTGTGCGCGGTATTTGCTTGAACAGGAGCGTGTGATGTTGGGCGAG	3780
QY	3781	ATCATCTTGTCCGGAAGCCGCTCATTTCCCGACAGGAGTCTTTACCGGGAGTTTC	3840
Db	3781	ATCATCTTGTCCGGAAGCCGCTCATTTCCCGACAGGAGTCTTTACCGGGAGTTTC	3840
QY	3841	GATGATGGAAGAGTGTGCGCTCACACCTCCCTTACATCGAAACAGGGAATGACGCTCGCC	3900
Db	3841	GATGATGGAAGAGTGTGCGCTCACACCTCCCTTACATCGAAACAGGGAATGACGCTCGCC	3900
QY	3901	GAAATTTTAAACAGAGGGAATCGGGTTGCTGCAAAACAGCCACCAAGCAAGCGGAGCT	3960
Db	3901	GAAATTTTAAACAGAGGGAATCGGGTTGCTGCAAAACAGCCACCAAGCAAGCGGAGCT	3960
QY	3961	GCTGCTCCGCTGTGGAATCCAAAGTGGCGGACCTCGAAGCGCTTCTGGCGGAAGCATATG	4020
Db	3961	GCTGCTCCGCTGTGGAATCCAAAGTGGCGGACCTCGAAGCGCTTCTGGCGGAAGCATATG	4020
QY	4021	TGGAATTTTCAATCAGCGGATACAATATTTAGAGGCTTGTCCACTCTGCTGCGCAACCC	4080
Db	4021	TGGAATTTTCAATCAGCGGATACAATATTTAGAGGCTTGTCCACTCTGCTGCGCAACCC	4080
QY	4081	CGATAGCATACATGATGGGATTCACAGCTCTATACAGCCGCTCACCACCCCAACAT	4140
Db	4081	CGATAGCATACATGATGGGATTCACAGCTCTATACAGCCGCTCACCACCCCAACAT	4140
QY	4141	ACCTCTCTGTTTAAATCCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCACGCGCT	4200
Db	4141	ACCTCTCTGTTTAAATCCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCACGCGCT	4200
QY	4201	GCTTCTGCTTTCTGATAGGCGCGGATCCTCGTGGAGCGGCTGTTGGCAGATAGGCCCTTGG	4260
Db	4201	GCTTCTGCTTTCTGATAGGCGCGGATCCTCGTGGAGCGGCTGTTGGCAGATAGGCCCTTGG	4260
QY	4261	BAAGTGTCTGTGGAATTTTGGCAGGTTATGGAGCAGGGTGGCAGGCGCGCTCGTGGCC	4320
Db	4261	BAAGTGTCTGTGGAATTTTGGCAGGTTATGGAGCAGGGTGGCAGGCGCGCTCGTGGCC	4320
QY	4321	TTTAAAGTCTATGAGCGCGGATGCCCTCCACCGAGGACCTGTTTAACTCTCCTGCT	4380
Db	4321	TTTAAAGTCTATGAGCGCGGATGCCCTCCACCGAGGACCTGTTTAACTCTCCTGCT	4380
QY	4381	ATCCTCTCCCTGGCGGCTAGTCTGGGGTCTGTGCGGAGCGATCTACTGCTGGGAC	4440
Db	4381	ATCCTCTCCCTGGCGGCTAGTCTGGGGTCTGTGCGGAGCGATCTACTGCTGGGAC	4440
QY	4441	GTGGCCCGCAGGGAGGGGCTGTGCAAGTGGATGAACCGCTCATAGCGTTCCGCTTCGCG	4500
Db	4441	GTGGCCCGCAGGGAGGGGCTGTGCAAGTGGATGAACCGCTCATAGCGTTCCGCTTCGCG	4500
QY	4501	GGTAACCACTCTCCCGCAGCACTATGTGCTGAGAGCGAGCGCTGACAGCGTGTCACT	4560
Db	4501	GGTAACCACTCTCCCGCAGCACTATGTGCTGAGAGCGAGCGCTGACAGCGTGTCACT	4560
QY	4561	CAGATCTCTCTAGTCTTACATCTAGCTGCTGAGAGGCTTCAACAGTGGATCAAC	4620
Db	4561	CAGATCTCTCTAGTCTTACATCTAGCTGCTGAGAGGCTTCAACAGTGGATCAAC	4620
QY	4621	GAGGACTCTCCACCGCATGCTCCGCTGCTGCTAAGAGATGTTTGGGATTTGATATGC	4680
Db	4621	GAGGACTCTCCACCGCATGCTCCGCTGCTGCTAAGAGATGTTTGGGATTTGATATGC	4680
QY	4681	ACGGTGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCCAGCTCCTGCGCGGATTC	4740
Db	4681	ACGGTGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCCAGCTCCTGCGCGGATTC	4740
QY	4741	GTCCCGCTTCTCTCATGTCAAGTGGGTACAGGGAGTCTGGGGGCGGACGCGATCATG	4800
Db	4741	GTCCCGCTTCTCTCATGTCAAGTGGGTACAGGGAGTCTGGGGGCGGACGCGATCATG	4800
QY	4801	CAAAACCACTGCTGAGCAGACATCAACGAGATGTCGAAATAACGGTTTCCATGAGG	4860
Db	4801	CAAAACCACTGCTGAGCAGACATCAACGAGATGTCGAAATAACGGTTTCCATGAGG	4860
QY	4861	ATCGTGGGCGCTTAGGACCTGTATGAACAGCTGGCATGAACATTCCTCCCATTAACGCGTAC	4920
Db	4861	ATCGTGGGCGCTTAGGACCTGTATGAACAGCTGGCATGAACATTCCTCCCATTAACGCGTAC	4920
QY	4921	ACACGGGCGCTTGCACGCGCTTCCCGGCGCAATTTCTTAGGCGCTGTGGCGGGTG	4980
Db	4921	ACACGGGCGCTTGCACGCGCTTCCCGGCGCAATTTCTTAGGCGCTGTGGCGGGTG	4980
QY	4981	GCTGCTGAGAGTACGTGGAGGTTACCGGGTGGGGATTTTCCACTACGTGACGGGCGATG	5040
Db	4981	GCTGCTGAGAGTACGTGGAGGTTACCGGGTGGGGATTTTCCACTACGTGACGGGCGATG	5040
QY	5041	ACCACTGACAAAGTAAAGTGGCGGCTGCTGAGGTTCCCGGCGCGGAAATTTCTTACAGAGTG	5100





```
QY 7261 ACTATTTCTTCTCCATCCTTCTAGCTCAGGAACACTTGAAAAAGCCCTAGATTGTGAG 7320
Db 7261 ACTATTTCTTCTCCATCCTTCTAGCTCAGGAACACTTGAAAAAGCCCTAGATTGTGAG 7320
QY 7321 ATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTACCTCAGATCATTCACGACTC 7380
Db 7321 ATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTACCTCAGATCATTCACGACTC 7380
QY 7381 CATGCCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440
Db 7381 CATGCCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440
QY 7441 TCATGCCCTCAGGAACACTTGGGTACCGCCCTTGCGAGTCTGGAGACATCGGGCCAGAAGT 7500
Db 7441 TCATGCCCTCAGGAACACTTGGGTACCGCCCTTGCGAGTCTGGAGACATCGGGCCAGAAGT 7500
QY 7501 GTCGCGCTAGGCTACTCTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560
Db 7501 GTCGCGCTAGGCTACTCTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560
QY 7561 AACTGGGCAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCGTCCCAAGTTGGAT 7620
Db 7561 AACTGGGCAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCGTCCCAAGTTGGAT 7620
QY 7621 TTATCCAGCTGGTTCTGTTCTGCTGTTACAGCGGGGAGACATATATACAGCCCTGTCTGT 7680
Db 7621 TTATCCAGCTGGTTCTGTTCTGCTGTTACAGCGGGGAGACATATATACAGCCCTGTCTGT 7680
QY 7681 GCCGACCCCGCTGTTCTGTTCTGCTGTTACCTTCTGTTAGGGTAGGCATCTAT 7740
Db 7681 GCCGACCCCGCTGTTCTGTTCTGCTGTTACCTTCTGTTAGGGTAGGCATCTAT 7740
QY 7741 CTACTCCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800
Db 7741 CTACTCCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800
QY 7801 TTCCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
Db 7801 TTCCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
QY 7861 TTTTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
Db 7861 TTTTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
QY 7921 TAGCTGTGAAGGTCGGTAGCGCTGTGACTCGAGAGTGCTGATCTGGCTCTCTGTC 7980
Db 7921 TAGCTGTGAAGGTCGGTAGCGCTGTGACTCGAGAGTGCTGATCTGGCTCTCTGTC 7980
QY 7981 AGATCAAGT 7989
Db 7981 AGATCAAGT 7989
```

## RESULT 6

```
AA147276
ID AAL47276 standard; DNA; 7992 BP.
XX
AC AAL47276;
XX
DT 30-AUG-2002 (first entry)
XX
DE Hepatitis C virus sub-genomic replicon clone I377-NS3-3'UTR.
XX
KW Hepatitis; HCV; core-neo; NS3 proteinase/helicase; vaccine; diagnosis;
KW virucide; hepatotropic; gene therapy; anti-viral; gene; ds.
XX
OS Hepatitis C virus.
XX
FH Key Location/Qualifiers
FT CDS 342..1181
FT /*tag= a
FT /product= "core-neo fusion protein"
```

```
FT CDS 1801..7758
FT /*tag= b
XX /product= "NS3 proteinase/helicase"
PN WO200238793-A2.
XX 16-MAY-2002.
XX
XX 02-NOV-2001; 2001WO-US046350.
XX
XX 07-NOV-2000; 2000US-0245866P.
XX (ANAD-) ANADYS PHARM INC.
XX
XX Bichko V;
XX
XX WPI; 2002-490082/52.
XX P-PSDB; AAO18000, AAO18001.
XX
XX Novel nucleic acid encoding replication competent recombinant hepatitis C
XX virus genome useful for screening anti-hepatitis C virus therapeutics and
XX for vaccine development.
XX
XX Claim 6; Page 43-47; 85pp; English.
XX
XX The present invention provides protein and coding sequences from
XX Hepatitis C virus (HCV), comprising all or part of the HCV genome and
XX able to replicate efficiently when transfected into a susceptible cell
XX line without reducing the growth rate of the cell line by more than 10
XX fold. The sequences are useful for screening for anti-HCV therapeutics,
XX for detecting antibodies to HCV in a biological sample such as blood,
XX serum, plasma, blood cells, lymphocytes, or liver cells from a subject,
XX for deriving authentic HCV components such as replication-complement non-
XX infectious, replication-defective infection-component, and replication-
XX defective non-infectious HCV, in gene therapy or gene vaccination
XX targeted to hepatic tissue for treating an animal infected or susceptible
XX to HCV infection and for studying HCV infection and propagation. The
XX present sequence is a clone of a fragment of the HCV genome which encodes
XX the core-neo and NS3 proteinase/helicase proteins
XX
XX Sequence 7992 BP; 1648 A; 2369 C; 2243 G; 1732 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 7987.4; DB 6; Length 7992;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GCCAGCCCCGATTTGGGGCGGACACTCCACATAGATCACTCCCTGTGAGGAACACTCTG 60
Db 1 GCCAGCCCCGATTTGGGGCGGACACTCCACATAGATCACTCCCTGTGAGGAACACTCTG 60
QY 61 TCTTCAACGCAAGGCTCTAGCCATGGGTTAGTATGATGCTGTCGAGCCCTCCAGGAC 120
Db 61 TCTTCAACGCAAGGCTCTAGCCATGGGTTAGTATGATGCTGTCGAGCCCTCCAGGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCAG 180
Db 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTCGCCAG 180
QY 181 GACGACCGGGTCTCTTCTTGGATCAACCCGCTCAATGCTCGTGGATTTGGGGCTGCCCCC 240
Db 181 GACGACCGGGTCTCTTCTTGGATCAACCCGCTCAATGCTCGTGGATTTGGGGCTGCCCCC 240
QY 241 CGGAGACTGTAGCCGAGTAGTGGTTCGGAAGGCTTGTGTACTGCTGCTGATAGG 300
Db 241 CGGAGACTGTAGCCGAGTAGTGGTTCGGAAGGCTTGTGTACTGCTGCTGATAGG 300
QY 301 GTGCTTGGAGTGCCCGGGAGGTCTCGTAGACCGGTGCAACCATGAGCAGCAATCTTAAAC 360
Db 301 GTGCTTGGAGTGCCCGGGAGGTCTCGTAGACCGGTGCAACCATGAGCAGCAATCTTAAAC 360
QY 361 CTCAAGAAAAACCAAGGGCGGCCATGATTGAACAAGATGGATTGACGAGGTTCTC 420
Db 361 CTCAAGAAAAACCAAGGGCGGCCATGATTGAACAAGATGGATTGACGAGGTTCTC 420
```

QY 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCAACAAGCAATCGGCTGCT 480  
DB 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCAACAAGCAATCGGCTGCT 480  
QY 481 CTGATGCGCGCTGTTCCGGCTGTGAGCGAGGGGCGCCGGTCTTTTGTTCGAAGACCG 540  
DB 481 CTGATGCGCGCTGTTCCGGCTGTGAGCGAGGGGCGCCGGTCTTTTGTTCGAAGACCG 540  
QY 541 ACCTGTCCGGTGCCTGAATGAATGAGGACGAGGAGCGCGCTATCGTGGCTGCGCA 600  
DB 541 ACCTGTCCGGTGCCTGAATGAATGAGGACGAGGAGCGCGCTATCGTGGCTGCGCA 600  
QY 601 CGAGCGGCTTCCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGAGGAGCTGGC 660  
DB 601 CGAGCGGCTTCCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGAGGAGCTGGC 660  
QY 661 TGTATTGGGCGAAGTCCCGGGGAGGATCTCTGTGATCTCACTTGTCTCCTGCCGAGA 720  
DB 661 TGTATTGGGCGAAGTCCCGGGGAGGATCTCTGTGATCTCACTTGTCTCCTGCCGAGA 720  
QY 721 AAGTATCCATATGCTGATGCAATGGGCGGCTGCAATAGCTTGATCCGGCTACCTGCC 780  
DB 721 AAGTATCCATATGCTGATGCAATGGGCGGCTGCAATAGCTTGATCCGGCTACCTGCC 780  
QY 781 CATTCGACCAACGAAGCAATCGCATCGAGCGAGCACTACTCTCGATGGAGCCGGTC 840  
DB 781 CATTCGACCAACGAAGCAATCGCATCGAGCGAGCACTACTCTCGATGGAGCCGGTC 840  
QY 841 TTGTCTGATCAGGATGATCTGGAAGAGGATCAGGGGCTCGGCCAGCGCACTGTTCG 900  
DB 841 TTGTCTGATCAGGATGATCTGGAAGAGGATCAGGGGCTCGGCCAGCGCACTGTTCG 900  
QY 901 CGAGGCTCAAGGCGCGATCCCGAAGAGGATCTCGTCTGAGCCCATGGCGATCCCT 960  
DB 901 CGAGGCTCAAGGCGCGATCCCGAAGAGGATCTCGTCTGAGCCCATGGCGATCCCT 960  
QY 961 GCTTTCGGAATATCATGCTGGAAATGGCGCTTTCTGGAATCATCGACTGGCGCGC 1020  
DB 961 GCTTTCGGAATATCATGCTGGAAATGGCGCTTTCTGGAATCATCGACTGGCGCGC 1020  
QY 1021 TGGGTGTGGCGGACCGCTATCAGACATAGCTTGGCTACCGGTGATTTGCTGAAGAGC 1080  
DB 1021 TGGGTGTGGCGGACCGCTATCAGACATAGCTTGGCTACCGGTGATTTGCTGAAGAGC 1080  
QY 1081 TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATGCGCGCTCCCGATTCGC 1140  
DB 1081 TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATGCGCGCTCCCGATTCGC 1140  
QY 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTCTCTGAGTTAAACAGACCAACG 1200  
DB 1141 AGCGCATCGCTTCTATCGCTTCTTACGAGTCTCTGAGTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCCCGCTAACGTTACTGGC 1260  
DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCGCCCGCTAACGTTACTGGC 1260  
QY 1261 CGAAGCGCTTGAATGAAGCGGCTGCGCTTCTATATGTTATTTCCACCATATTG 1320  
DB 1261 CGAAGCGCTTGAATGAAGCGGCTGCGCTTCTATATGTTATTTCCACCATATTG 1320  
QY 1321 CCGTCTTTTGGCAATGAGGGCCCGGAAACCTGGCCCTCTCTTGTGACGACATTCCT 1380  
DB 1321 CCGTCTTTTGGCAATGAGGGCCCGGAAACCTGGCCCTCTCTTGTGACGACATTCCT 1380  
QY 1381 AGGGGCTTTTCCCTCTCGCGCAAGGAATGCAAGGCTGTGTTGAATGTCGTGAAGGAGCA 1440  
DB 1381 AGGGGCTTTTCCCTCTCGCGCAAGGAATGCAAGGCTGTGTTGAATGTCGTGAAGGAGCA 1440  
QY 1441 GTTTCCTGGAAGCTTTCTGAAGACAAACAACTGTGTAGCGACCTTTTGCAGGACGCG 1500  
DB 1441 GTTTCCTGGAAGCTTTCTGAAGACAAACAACTGTGTAGCGACCTTTTGCAGGACGCG 1500

QY 1501 AACCCCCCACTGCGGACAGGTCCTCTGCGGCCAAAGCCACCGTGTATAGATACACT 1560  
DB 1501 AACCCCCCACTGCGGACAGGTCCTCTGCGGCCAAAGCCACCGTGTATAGATACACT 1560  
QY 1561 GCAAAGCGGCACAAACCCAGTGCACGTTGTGTGAGTTGGATAGTTGTGGAAGAGTCAAA 1620  
DB 1561 GCAAAGCGGCACAAACCCAGTGCACGTTGTGTGAGTTGGATAGTTGTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAAGAGTACCCCATTTG 1680  
DB 1621 TGGCTCTCTCAAGCGTATTCAACAAGGGGCTGAAGGATGCCAAGAGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGGCTCTGGTGACATGCTTTTACATGTTTGTAGTTCGAGGTTAAA 1740  
DB 1681 ATGGGATCTGATCTGGGGCTCTGGTGACATGCTTTTACATGTTTGTAGTTCGAGGTTAAA 1740  
QY 1741 AACGCTTAGGCCCCCGCAACACAGGGGACCTGGTTTTTCTTTGAAAACACGATAATACC 1800  
DB 1741 AACGCTTAGGCCCCCGCAACACAGGGGACCTGGTTTTTCTTTGAAAACACGATAATACC 1800  
QY 1801 ATGGCGCTATTACGGCTACTCCCAACAGACGAGGCGCTACTTTGGCTGCATCATCACT 1860  
DB 1801 ATGGCGCTATTACGGCTACTCCCAACAGACGAGGCGCTACTTTGGCTGCATCATCACT 1860  
QY 1861 AGCCTCACAGCGCGGACAGGAACAGGTCGAGGGGAGGTCGAAGTGTCTCCACCGCA 1920  
DB 1861 AGCCTCACAGCGCGGACAGGAACAGGTCGAGGGGAGGTCGAAGTGTCTCCACCGCA 1920  
QY 1921 ACACATCTTTCTTGGCGACTCGTCAATGGCGTGTGTTGGACTGCTCATCATGTGCC 1980  
DB 1921 ACACATCTTTCTTGGCGACTCGTCAATGGCGTGTGTTGGACTGCTCATCATGTGCC 1980  
QY 1981 GGCTCAAGACCCCTTGCGCGCCCAAGGGGCCCAATCACCCAAATGTACACCAATGTGGAC 2040  
DB 1981 GGCTCAAGACCCCTTGCGCGCCCAAGGGGCCCAATCACCCAAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACCTCGTCCGCTGGCAAGCGCCCCCGGGCGGCTTCTTTGACACCATGCACTGCG 2100  
DB 2041 CAGGACCTCGTCCGCTGGCAAGCGCCCCCGGGCGGCTTCTTTGACACCATGCACTGCG 2100  
QY 2101 GGCACTCGGACCTTTACTTGTTCAGAGGATGCGGATGTCATTCGGTGGCGCGCGG 2160  
DB 2101 GGCACTCGGACCTTTACTTGTTCAGAGGATGCGGATGTCATTCGGTGGCGCGCGG 2160  
QY 2161 GGCGACAGAGGGGAGCTTCTCTCCCGAGCGCGTCTCTTACTTGAAGGCTCTTCG 2220  
DB 2161 GGCGACAGAGGGGAGCTTCTCTCCCGAGCGCGTCTCTTACTTGAAGGCTCTTCG 2220  
QY 2221 GCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGATCTTTTGGGCTGCGGTGTC 2280  
DB 2221 GCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGATCTTTTGGGCTGCGGTGTC 2280  
QY 2281 ACCCGAGGGGTTCCGAAGCGGTGGACTTTGTACCGTTCGAGTCTATGGAACCACTATG 2340  
DB 2281 ACCCGAGGGGTTCCGAAGCGGTGGACTTTGTACCGTTCGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCCGCTTCTACGACAACTCGTCCCTCCCGCGCTACCGCAGACATTTCCAGGTG 2400  
DB 2341 CGGTCCCCGCTTCTACGACAACTCGTCCCTCCCGCGCTACCGCAGACATTTCCAGGTG 2400  
QY 2401 GCCCATCTACAGCCCTTACTGTAGCGGCAAGACACTAAGTGCCTGCGGTGCGTATGCA 2460  
DB 2401 GCCCATCTACAGCCCTTACTGTAGCGGCAAGACACTAAGTGCCTGCGGTGCGTATGCA 2460  
QY 2461 GCCCAAGGTATAAGGTGCTTGTCTGAAACCGCTCGCTCGCGCCACCTTAGGTTTCGGG 2520  
DB 2461 GCCCAAGGTATAAGGTGCTTGTCTGAAACCGCTCGCTCGCGCCACCTTAGGTTTCGGG 2520  
QY 2521 GGTATATGCTAAGGCACATGATCGACCTTAAACATCAGAACCGGGGTAAAGACCATC 2580  
DB 2521 GGTATATGCTAAGGCACATGATCGACCTTAAACATCAGAACCGGGGTAAAGACCATC 2580  
QY 2581 ACCAGGGGTGCCCCCATACGTACTCCACTATGGCAAGTTTCTTGGCGAGCGGTGCTGC 2640



Qy	4801	CAAA	CCACTCTCCCATGTGGAGCA	CAGATCA	CCGACATGTGA	AAACCGTGTCCATG	GCG	4860
Db	4801	CAB	ACCACCTGCGCATGTGGAGCA	CAGATCA	CCGACATGTGA	AAACCGTGTCCATG	GCG	4860
Qy	4861	ATCG	TGGGCGCTAGGACCTGTAGTA	ACAGCTGTG	GAACATTTCCCATTA	ACCGCTAC	4920	
Db	4861	ATCG	TGGGCGCTAGGACCTGTAGTA	ACAGCTGTG	GAACATTTCCCATTA	ACCGCTAC	4920	
Qy	4921	ACCA	GGGCCCCCTGCACGCCCTCCCGCGC	CAAAATTTCTAGGGCGCTGTGCGCGGT	4980			
Db	4921	ACC	GGGCCCCCTGCACGCCCTCCCGCGC	CAAAATTTCTAGGGCGCTGTGCGCGGT	4980			
Qy	4981	GCT	GCTGAGGAGTACGTGGAGTTAC	CGCGGTGGGGATTTCCACTACGTAC	CGGCATG	5040		
Db	4981	GCT	GCTGAGGAGTACGTGGAGTTAC	CGCGGTGGGGATTTCCACTACGTAC	CGGCATG	5040		
Qy	5041	ACCA	CTGACAACTGTAAGTCCCGTGT	CAGGTTC	CGGCCCGCGAAATTTCTTCA	CAGAACTG	5100	
Db	5041	ACCA	CTGACAACTGTAAGTCCCGTGT	CAGGTTC	CGGCCCGCGAAATTTCTTCA	CAGAACTG	5100	
Qy	5101	GAT	GGGTGCGGTGCACAGTACGCT	CCACGCTGCAAA	CCCCCTCTACGGGAGAGT	5160		
Db	5101	GAT	GGGTGCGGTGCACAGTACGCT	CCACGCTGCAAA	CCCCCTCTACGGGAGAGT	5160		
Qy	5161	ACA	TTCTGTGCGGCTCAATCAAT	ACTGTGTGGGTACAGT	TCCCATCGCAGGCCGAA	5220		
Db	5161	ACA	TTCTGTGCGGCTCAATCAAT	ACTGTGTGGGTACAGT	TCCCATCGCAGGCCGAA	5220		
Qy	5221	CGG	ACGTAGAGTGCTCACTTCA	TGCTCA	CGACCCCTCCACATTCACGGCGGAG	5280		
Db	5221	CGG	ACGTAGAGTGCTCACTTCA	TGCTCA	CGACCCCTCCACATTCACGGCGGAG	5280		
Qy	5281	GCT	AAAGCTAGGCTGGCCACGGGAT	CTCCCGCTCTTGGCGAGCT	CATCAGCTAGCCAG	5340		
Db	5281	GCT	AAAGCTAGGCTGGCCACGGGAT	CTCCCGCTCTTGGCGAGCT	CATCAGCTAGCCAG	5340		
Qy	5341	CTG	CTGCGCTCTCTTTGAAGGCA	ATGCATCACCTACCGT	CTACGCTCCCGGACGCTGAC	5400		
Db	5341	CTG	CTGCGCTCTCTTTGAAGGCA	ATGCATCACCTACCGT	CTACGCTCCCGGACGCTGAC	5400		
Qy	5401	CTCA	TGAGGCCAACTCTCTGTGCGGC	ACGAGATGGCGGGAACTCA	CCCGCTGGAG	5460		
Db	5401	CTCA	TGAGGCCAACTCTCTGTGCGGC	ACGAGATGGCGGGAACTCA	CCCGCTGGAG	5460		
Qy	5461	TCAGA	AAATAAAGTAGTAGTAATTTTGG	ACTCTTCGACCGCTTCAAGGGAGGAGTAG	5520			
Db	5461	TCAGA	AAATAAAGTAGTAGTAATTTTGG	ACTCTTCGACCGCTTCAAGGGAGGAGTAG	5520			
Qy	5521	AGG	AAAGTATCCGTTCCGGCGGAGAT	CTCTGCGGAGTCCAGGAAATTTCCCTCGACGATG	5580			
Db	5521	AGG	AAAGTATCCGTTCCGGCGGAGAT	CTCTGCGGAGTCCAGGAAATTTCCCTCGACGATG	5580			
Qy	5581	CCCA	TATGGGCACGCGCGGATTA	CAACCTTCACTGTTAGAGTCTTGAAGGAC	CCCGAC	5640		
Db	5581	CCCA	TATGGGCACGCGCGGATTA	CAACCTTCACTGTTAGAGTCTTGAAGGAC	CCCGAC	5640		
Qy	5641	TAC	GTCCCTCCAGTGTATACAGG	GTTCATTTGCGCTTCAAGCGCTTCCGATACCA	5700			
Db	5641	TAC	GTCCCTCCAGTGTATACAGG	GTTCATTTGCGCTTCAAGCGCTTCCGATACCA	5700			
Qy	5701	CCT	CCA	CGGAGGAGGACGTTGCTGT	CTCAGAAATCTACCGTGTCTTCTGCTTGGG	5760		
Db	5701	CCT	CCA	CGGAGGAGGAGGACGTTGCTGT	CTCAGAAATCTACCGTGTCTTCTGCTTGGG	5760		
Qy	5761	GAG	CTGCCACAAGACCTTCGG	CAGCTCCGAATCTGTGCGCGTTCGACAGCGGCA	5820			
Db	5761	GAG	CTGCCACAAGACCTTCGG	CAGCTCCGAATCTGTGCGCGTTCGACAGCGGCA	5820			
Qy	5821	ACG	GCCTCTCTCACCAGCCCTCCG	ACGACGCGGATCCGACGTTGAGTCGTAC	5880			
Db	5821	ACG	GCCTCTCTCACCAGCCCTCCG	ACGACGCGGATCCGACGTTGAGTCGTAC	5880			

QY	5881	TCTCTCATGCCCCCTTTGAGGGGAGCCGGGGATCCGATCTACGACGGGTCTTGG	594
Db	5881	TCTCTCATGCCCCCTTTGAGGGGAGCCGGGGATCCGATCTACGACGGGTCTTGG	5940
QY	5941	TCCTACCGTAAAGCGAGGAGCTAGTGAAGACGTCGCTCTGCTCTCGATGTCTTACATGG	6000
Db	5941	TCCTACCGTAAAGCGAGGAGCTAGTGAAGACGTCGCTCTGCTCTCGATGTCTTACATGG	6000
QY	6001	ACAGCGCCCTGATCACGCCATCGCTCGCGAGGAAACCAAGCTGCCCATCAATGCACTG	6060
Db	6001	ACAGCGCCCTGATCACGCCATCGCTCGCGAGGAAACCAAGCTGCCCATCAATGCACTG	6060
QY	6061	AGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAACTCTCGACGCGAAGC	6120
Db	6061	AGCAACTCTTTGCTCCGTCACCAAACTTGGTCTATGCTACAACTCTCGACGCGAAGC	6120
QY	6121	CTGCGGCAAGAAGGTACACTTTGACAGACTCGCAGTCTGGAAGACCACTACCGGGAC	6180
Db	6121	CTGCGGCAAGAAGGTACACTTTGACAGACTCGCAGTCTGGAAGACCACTACCGGGAC	6180
QY	6181	GTGCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTAAAGGCTAAACTTCTATCCGTGGAG	6240
Db	6181	GTGCTCAAGGAGATGAAGGCGAAGCGCTCCACAGTTAAAGGCTAAACTTCTATCCGTGGAG	6240
QY	6241	GAAGCTGTAAAGTGAAGGCGAAGCGCTCCACAGTTAAAGGCTAAACTTCTATCCGTGGAG	6300
Db	6241	GAAGCTGTAAAGTGAAGGCGAAGCGCTCCACAGTTAAAGGCTAAACTTCTATCCGTGGAG	6300
QY	6301	GACGTCGGAACCTATCCAGCAAGCGGTTAAACCACTCCGCTCCGCTGGAAGGACTTGG	6360
Db	6301	GACGTCGGAACCTATCCAGCAAGCGGTTAAACCACTCCGCTCCGCTGGAAGGACTTGG	6360
QY	6361	CTGGAAGCACTGAGACACCAATTTGACACCAACCATCATGGCAAAATAGGTTTCTGC	6420
Db	6361	CTGGAAGCACTGAGACACCAATTTGACACCAACCATCATGGCAAAATAGGTTTCTGC	6420
QY	6421	GTCACACGAGAAAGGGGGCGGCAAGCTCGCTTATCGTATTCACAGATTGGGG	6480
Db	6421	GTCACACGAGAAAGGGGGCGGCAAGCTCGCTTATCGTATTCACAGATTGGGG	6480
QY	6481	GTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCCGTG	6540
Db	6481	GTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCCCTCAGGCCGTG	6540
QY	6541	ATGGGCTCTTCTACCGATTCCAAATACTCTCTCGGACAGCGGTGAGTCTCTGTGTAAT	6600
Db	6541	ATGGGCTCTTCTACCGATTCCAAATACTCTCTCGGACAGCGGTGAGTCTCTGTGTAAT	6600
QY	6601	GCCTGGAAGCGGAATAATGCCCTATGGGCTTGCATATGACACCGCTGTTTGACTCA	6660
Db	6601	GCCTGGAAGCGGAATAATGCCCTATGGGCTTGCATATGACACCGCTGTTTGACTCA	6660
QY	6661	ACGCTCACTCAGAAATGACATCCGTTGTAGGAGTCAATCTACCAATGTGTGACTTGGCC	6720
Db	6661	ACGCTCACTCAGAAATGACATCCGTTGTAGGAGTCAATCTACCAATGTGTGACTTGGCC	6720
QY	6721	CCGGAAGCCAGACAGGCCATTAAGTTCGCTCACAGAGCGCTTTACATCGGGGGCCCCCTG	6780
Db	6721	CCGGAAGCCAGACAGGCCATTAAGTTCGCTCACAGAGCGCTTTACATCGGGGGCCCCCTG	6780
QY	6781	ACTAATTTTAAAGGCGAGAACTGCGGCTATCCCGGTCCCGCGAGCGGTGTACTGACG	6840
Db	6781	ACTAATTTTAAAGGCGAGAACTGCGGCTATCCCGGTCCCGCGAGCGGTGTACTGACG	6840
QY	6841	ACCAGCTCGGTAAATACCTCTCATGTTACTTTGAAGGCGGCTGCGGCTGTGAGCTGCG	6900
Db	6841	ACCAGCTCGGTAAATACCTCTCATGTTACTTTGAAGGCGGCTGCGGCTGTGAGCTGCG	6900
QY	6901	AAGCTCCAGACTTGCACGATGCTCGTATGCGGAGACGACCTTGTGCTTATCTGTGAAGC	6960
Db	6901	AAGCTCCAGACTTGCACGATGCTCGTATGCGGAGACGACCTTGTGCTTATCTGTGAAGC	6960
QY	6961	GCGGGACCCAGAGGAGCGAGCCTTACGCGGCTTACGCGGCTTACGCTAGATATAC	7020

Db 6961 GCGGGGACCCAGAGGACGAGGCGAGCCTACGGGCTTCACGGAGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCGCCCTCGGGACCCGCCCAACCAAGATACGACTTGGAGTTGATAACATCATGC 7080  
Db 7021 TCTGCCGCCCTCGGGACCCGCCCAACCAAGATACGACTTGGAGTTGATAACATCATGC 7080  
QY 7081 TCCTCCAAATGTGTCAGTGCAGCAGATCATCTGCGCAAAAGGGTGATATCTCACCCGT 7140  
Db 7081 TCCTCCAAATGTGTCAGTGCAGCAGATCATCTGCGCAAAAGGGTGATATCTCACCCGT 7140  
QY 7141 GACCCACACCCGCCCTCGGGGCTGCTGGGAGACAGTACACACATCCAGTCAAT 7200  
Db 7141 GACCCACACCCGCCCTCGGGGCTGCTGGGAGACAGTACACACATCCAGTCAAT 7200  
QY 7201 TCCTGGCTAGGCAACATCATATGATGCGGCCACCTTGTGGCAAGGATGATCTGATG 7260  
Db 7201 TCCTGGCTAGGCAACATCATATGATGCGGCCACCTTGTGGCAAGGATGATCTGATG 7260  
QY 7261 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGAAGAGCCCTAGATTGTCTAG 7320  
Db 7261 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGAAGAGCCCTAGATTGTCTAG 7320  
QY 7321 ATCTACGGGGCTGTACTCCATTCAGCCACTTGCACCTACCTCAGATCATTCACAGCTC 7380  
Db 7321 ATCTACGGGGCTGTACTCCATTCAGCCACTTGCACCTACCTCAGATCATTCACAGCTC 7380  
QY 7381 CATGGCTTAGCGCACTTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGT 7440  
Db 7381 CATGGCTTAGCGCACTTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGT 7440  
QY 7441 TCATGGCTCAGAAACTTGGGGTACCGCCCTTGCAGTCTGAGACATCGGCCAGAGT 7500  
Db 7441 TCATGGCTCAGAAACTTGGGGTACCGCCCTTGCAGTCTGAGACATCGGCCAGAGT 7500  
QY 7501 GTCCGGCTAGGCTACTGTCACAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Db 7501 GTCCGGCTAGGCTACTGTCACAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
QY 7561 AACTGGGCACTAAGGACCAAGCTCAAACTCACTCCAAATCCCGCTGCGTCCAGTTGGAT 7620  
Db 7561 AACTGGGCACTAAGGACCAAGCTCAAACTCACTCCAAATCCCGCTGCGTCCAGTTGGAT 7620  
QY 7621 TTATCAGCTGGTTGCTGCTGTGTACAGGGGGAGACATATATCAGACCTGTCTGCT 7680  
Db 7621 TTATCAGCTGGTTGCTGCTGTGTACAGGGGGAGACATATATCAGACCTGTCTGCT 7680  
QY 7681 GCGCGACCCGCTGGTTGCTGTGTGCTGCTACTCTCTACTTTCTGTAGGGGTAGGCATCTAT 7740  
Db 7681 GCGCGACCCGCTGGTTGCTGTGTGCTGCTACTCTCTACTTTCTGTAGGGGTAGGCATCTAT 7740  
QY 7741 CTACTCCCAACCGATGAAGGGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTT 7800  
Db 7741 CTACTCCCAACCGATGAAGGGGACCTAAACACTCCAGGCCAATAGGCCATCTGTTT 7800  
QY 7801 TTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db 7801 TTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
QY 7861 TTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db 7861 TTTCCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
QY 7921 TAGCTGTGAAGCTCCGTGAGCCGCTGTACTGAGAGAGTGTGATGCTGGCCTCTCTGC 7980  
Db 7921 TAGCTGTGAAGCTCCGTGAGCCGCTGTACTGAGAGAGTGTGATGCTGGCCTCTCTGC 7980  
QY 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989

ABK91412  
ID ABK91412 standard; DNA; 10690 BP.  
XX  
AC ABK91412;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus vector construct pHCVNeo.17.  
XX  
KW HCV; ss; pHCVNeo.17; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRBS; NS5A; HCV replication.  
XX  
OS Hepatitis C virus.  
OS Encephalomyocarditis virus.  
OS Escherichia coli.  
OS Enterobacteria phage T7.  
OS Synthetic.  
XX  
Key Location/Qualifiers  
5'UTR 1..341  
/\*tag= a  
CDS 342..1181  
/\*tag= b  
/\*product= "Core-neo fusion protein"  
misc\_signal 1190..1800  
/\*tag= c  
/\*label= IRES  
/\*note= "Internal ribosome entry site from EMCV"  
CDS 1801..7758  
/\*tag= d  
/\*product= "Polyprotein"  
/\*note= "Comprising NS3, NS4A, NS4B, NS5A and NS5B"  
3'UTR 7759..7989  
/\*tag= e  
misc\_feature 7990..10690  
/\*tag= f  
/\*note= "Plasmid derived sequences"  
WO200259321-A2.  
01-AUG-2002.  
16-JAN-2002; 2002WO-EP000526.  
23-JAN-2001; 2001US-0263479P.  
(RICE-) IST RICERCHE BIOL MOLECOLARE ANGELETTI.  
De Francesco R, Migliaccio G, Paonessa G;  
WPI; 2002-599793/64.  
XX  
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
XX ribosome entry site (IRES) region, useful in studying HCV replication and  
XX expression.  
XX Claim 1; Fig 1; 69pp; English.  
XX The invention relates to nucleic acid molecules comprising altered HCV  
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)  
XX internal ribosome entry site (IRES) region coding for one or more NS3,  
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations  
XX are detailed in the specification. Also included are (1) an expression  
XX vector comprising a nucleotide sequence coding for the altered nucleic  
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a  
XX recombinant cell human hepatoma cell produced by introducing into a human hepatoma  
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)  
XX replicon enhanced cell or which containing a functional HCV replicon; (5)  
XX an HCV replicon enhanced cells made in the method; and (6) measuring the  
XX ability of a compound to affect HCV activity. The HCV replicons and HCV



CC replicon enhanced cells are useful in studying HCV replication and  
CC expression, and HCV and host cell interactions, producing HCV RNA and  
CC proteins, and providing a system for measuring the ability of a compound  
CC to modulate one or more HCV activities e.g. to discover drugs which may  
CC treat HCV mediated diseases such as liver failure, cirrhosis and  
CC hepatocellular carcinoma. The present sequence is the HCV based vector  
CC pHCVneo.17, used as a basis for the adaptive mutations of the invention  
XX

SQ Sequence 10690 BP; 2334 A; 3045 C; 2908 G; 2403 T; 0 U; 0 Other;

Query Match 100.0%; Score 7987.4; DB 6; Length 10690;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY	1	GCCAGCCCCGATGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTAGT	60
DB	1	GCCAGCCCCGATGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTAGT	60
QY	61	TCCTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTCAGCCTCCAGGAC	120
DB	61	TCCTCAGCAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTCTGTCAGCCTCCAGGAC	120
QY	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTGCCAG	180
DB	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTGCCAG	180
QY	181	GAGACCGGGTCTCTTCTTGATCAACCCGCTCAATGCTGGAGATTGGCGGTGCCCC	240
DB	181	GAGACCGGGTCTCTTCTTGATCAACCCGCTCAATGCTGGAGATTGGCGGTGCCCC	240
QY	241	GCGAGACTGTAGCCGAGTAGTGTGGTCCGAAAGCGCTTGTGCTACTGCTGTATAGG	300
DB	241	GCGAGACTGTAGCCGAGTAGTGTGGTCCGAAAGCGCTTGTGCTACTGCTGTATAGG	300
QY	301	GTCTTTCGAGTGGCCCGGAGGTCTCTGAGACCGTGCACCATGAGCAAGAACTCAAAC	360
DB	301	GTCTTTCGAGTGGCCCGGAGGTCTCTGAGACCGTGCACCATGAGCAAGAACTCAAAC	360
QY	361	CTCAAGAAACCAAGAGCGCGCCATGATTGAACAGATGATTCACGCGAGTTCTC	420
DB	361	CTCAAGAAACCAAGAGCGCGCCATGATTGAACAGATGATTCACGCGAGTTCTC	420
QY	421	CGCGCGCTGGGTGGAGAGCTATTGGCTATGACTGGGCAACACAGCAATCGGTGCT	480
DB	421	CGCGCGCTGGGTGGAGAGCTATTGGCTATGACTGGGCAACACAGCAATCGGTGCT	480
QY	481	CTGATCGCCGCTGTTCCGGCTGTGAGCGCAGGGGCGCCGGTCTTTTGTCAAGACCG	540
DB	481	CTGATCGCCGCTGTTCCGGCTGTGAGCGCAGGGGCGCCGGTCTTTTGTCAAGACCG	540
QY	541	ACCTGTCCGCTGCCCTGAAATGAATGAACTGACGAGCAGGCGCGGCTATCGTGGCTGCCA	600
DB	541	ACCTGTCCGCTGCCCTGAAATGAATGAACTGACGAGCAGGCGCGGCTATCGTGGCTGCCA	600
QY	601	CGACGGGCTTCTTTCGCGAGCTGTCTCGACGTTGCTACTGAAGCGGAAGGAGCTGGC	660
DB	601	CGACGGGCTTCTTTCGCGAGCTGTCTCGACGTTGCTACTGAAGCGGAAGGAGCTGGC	660
QY	661	TGCTATTGGGCAAGTGGCGGGGAGGATCTCTGTCTCTCACTTGTCTTCTGCGGAGA	720
DB	661	TGCTATTGGGCAAGTGGCGGGGAGGATCTCTGTCTCTCACTTGTCTTCTGCGGAGA	720
QY	721	AAGTATCCATCATGGCTGATGAATGCAATGCGGCGGCTGATACGTTGCTTCTGCGGAGA	780
DB	721	AAGTATCCATCATGGCTGATGAATGCAATGCGGCGGCTGATACGTTGCTTCTGCGGAGA	780
QY	781	CATTGACCAACGAGCAACATCGCATCGAGCGACGCTACTCGGATGGAGCGGCTC	840
DB	781	CATTGACCAACGAGCAACATCGCATCGAGCGACGCTACTCGGATGGAGCGGCTC	840
QY	841	TTGTGATCAGATGATCTGAGCGAAGCATCAGGGGCTCGCGCAGCGAAGTGTTCG	900
DB	841	TTGTGATCAGATGATCTGAGCGAAGCATCAGGGGCTCGCGCAGCGAAGTGTTCG	900

QY	901	CCAGGCTCAAGCGCGCATGCCGACGCGAGGATCTCTGTGACCCATGCGGATGCT	960
DB	901	CCAGGCTCAAGCGCGCATGCCGACGCGAGGATCTCTGTGACCCATGCGGATGCT	960
QY	961	GCTTGCAGAAATATCATGTGTGAAATGCGCGCTTTCTGGAATTCATCGACTGTGCCGCGC	1020
DB	961	GCTTGCAGAAATATCATGTGTGAAATGCGCGCTTTCTGGAATTCATCGACTGTGCCGCGC	1020
QY	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCGTGGCTACCGGTGATATTCGTAAGAGC	1080
DB	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCGTGGCTACCGGTGATATTCGTAAGAGC	1080
QY	1081	TTGCGCGGAATGGGCTGACCGCTTCTCGTGTCTTACGCTATCGCGCTCCCGATTCCG	1140
DB	1081	TTGCGCGGAATGGGCTGACCGCTTCTCGTGTCTTACGCTATCGCGCTCCCGATTCCG	1140
QY	1141	AGCGCATCGCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACG	1200
DB	1141	AGCGCATCGCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACG	1200
QY	1201	GTTTCCCTTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAGCTTACTGGC	1260
DB	1201	GTTTCCCTTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAGCTTACTGGC	1260
QY	1261	CGAAGCGCTTGAATTAAGCGCGGTGTGCGCTTGTCTATATGTTTATTTCCACCATATG	1320
DB	1261	CGAAGCGCTTGAATTAAGCGCGGTGTGCGCTTGTCTATATGTTTATTTCCACCATATG	1320
QY	1321	CGCTCTTTTGGAAATGTGAGGCGCGGAAACCTTGGCCCTCTTCTTTCGAGAGATTCCT	1380
DB	1321	CGCTCTTTTGGAAATGTGAGGCGCGGAAACCTTGGCCCTCTTCTTTCGAGAGATTCCT	1380
QY	1381	AGGGCTCTTTCCTCTCGCCAAAGGAATCGAAGTCTGTGTGATGTCTGTGAGGAGACGA	1440
DB	1381	AGGGCTCTTTCCTCTCGCCAAAGGAATCGAAGTCTGTGTGATGTCTGTGAGGAGACGA	1440
QY	1441	GTTCTCTGGAAGCTTCTTGAAGACAAACAACTCTGTAGCGACCCCTTTCGAGGACGGG	1500
DB	1441	GTTCTCTGGAAGCTTCTTGAAGACAAACAACTCTGTAGCGACCCCTTTCGAGGACGGG	1500
QY	1501	AACCCCTCACCTGCGACAGGTGCTCTGCGGCAAAAGCCAGTGTATAAGATACACCT	1560
DB	1501	AACCCCTCACCTGCGACAGGTGCTCTGCGGCAAAAGCCAGTGTATAAGATACACCT	1560
QY	1561	GCAAGGCGGCAACAAACCCAGTGCACGTTGTGAGTTGATGTCTGTGAAAGTCAAA	1620
DB	1561	GCAAGGCGGCAACAAACCCAGTGCACGTTGTGAGTTGATGTCTGTGAAAGTCAAA	1620
QY	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT	1680
DB	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT	1680
QY	1681	ATGGGATCTGATCTGGGCTCGGTGCAATGCTTTCATGTTTGTGAGGTTTAAAA	1740
DB	1681	ATGGGATCTGATCTGGGCTCGGTGCAATGCTTTCATGTTTGTGAGGTTTAAAA	1740
QY	1741	AACGCTCTAGGCGCCCGGACCAACCGGAGCGTGGTCTTCTTGAAGAAACACCATATAC	1800
DB	1741	AACGCTCTAGGCGCCCGGACCAACCGGAGCGTGGTCTTCTTGAAGAAACACCATATAC	1800
QY	1801	ATGGCGCTTATTCAGCGCTTCTCCCAACAGCGAGGCGCTACTTGGCTGCAATCACT	1860
DB	1801	ATGGCGCTTATTCAGCGCTTCTCCCAACAGCGAGGCGCTACTTGGCTGCAATCACT	1860
QY	1861	AGCCTCAGAGCGCGGACAGGAACAGGTGAGGGGAGGTCCAAAGTGTCTTCCACCGCA	1920
DB	1861	AGCCTCAGAGCGCGGACAGGAACAGGTGAGGGGAGGTCCAAAGTGTCTTCCACCGCA	1920
QY	1921	ACCAATCTTCTTGGCGACCTGGTCAATGCGGTGTGTGAGTGTCTATCATGTGCTCC	1980
DB	1921	ACCAATCTTCTTGGCGACCTGGTCAATGCGGTGTGTGAGTGTCTATCATGTGCTCC	1980







```
FT mat_peptide 3697..3858
FT /*tag= f
FT /note= "Non-structural protein 4A"
FT mat_peptide 3859..4641
FT /*tag= g
FT /note= "Non-structural protein 4B"
FT mat_peptide 4642..5982
FT /*tag= h
FT /note= "Non-structural protein 5A"
FT mat_peptide 5983..7755
FT /*tag= i
FT /note= "Non-structural protein 5B and RNA-dependent RNA
FT polymerase"
FT 3' UTR 7759..7989
FT /*tag= j
FT /note= "HCV 3'UTR"
XX
XX US2003045568-A1.
XX
XX 06-MAR-2003.
XX
XX 19-APR-2002; 2002US-00125940.
XX
XX 20-APR-2001; 2001US-0285104P.
XX
XX (ALTA//) ALTAMURA S.
XX (KOCH//) KOCH U.
XX
XX Altamura S, Koch U;
XX WPI; 2003-447715/42.
XX
XX Treating infections by hepatitis C virus and its related conditions
XX comprises administering thiosemicarbazone compounds.
XX
XX Example 2; Page 9-14; 25pp; English.
XX
XX The invention relates to a method of treating or preventing infection by
XX hepatitis C virus or its related conditions by delaying the onset and
XX inhibiting replication of hepatitis C virus which comprises administering
XX thiosemicarbazone compounds. The method is useful for treating or
XX preventing infection by hepatitis C virus or its related conditions e.g.
XX liver inflammation, liver failure or cirrhosis, delaying the onset and
XX inhibiting replication of hepatitis C. The present sequence represents
XX the hepatitis C virus expression plasmid pHCVNeo17.wt DNA
XX
XX Sequence 10690 BP; 2334 A; 3045 C; 2908 G; 2403 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 7987.4; DB 9; Length 10690;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
XX
XX 1 GCCAGCCCCCGATTGGGGGACACTCCACATAGATCACTCCCTGTGAGGAACACTG 60
XX
XX 1 GCCAGCCCCCGATTGGGGGACACTCCACATAGATCACTCCCTGTGAGGAACACTG 60
XX
XX 61 TCTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGTCGACCTCCAGGAC 120
XX
XX 61 TCTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGTCGACCTCCAGGAC 120
XX
XX 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGAACCGGTGAGTACACCGGAATGCCAG 180
XX
XX 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGAACCGGTGAGTACACCGGAATGCCAG 180
XX
XX 181 GAGACCGGGTCTTCTTGGATCAACCGCTCAATGCTGTGAGATTGGGGTGCCCCC 240
XX
XX 181 GAGACCGGGTCTTCTTGGATCAACCGCTCAATGCTGTGAGATTGGGGTGCCCCC 240
XX
XX 241 GCGAGACTCTAGCCGAGTAGTGTGGGTGCGAAAGGCGCTTGTGCTACTGCTGATAGG 300
XX
XX 241 GCGAGACTCTAGCCGAGTAGTGTGGGTGCGAAAGGCGCTTGTGCTACTGCTGATAGG 300
XX
XX 301 GTGCTTGCGAGTGCCCCGGAGGCTCTCGTAGACCGGTGACCATGAGCACGAATCCTAAAC 360
```

```
Db 301 GTGCTTGCGAGTGCCCCGGAGGCTCTCGTAGACCGGTGACCATGAGCACGAATCCTAAAC 360
QY 361 CTCAAAGAAAACCAAGGGCGGCCATGATTGAAACAAGATGTCACGAGGTTCCTC 420
Db 361 CTCAAAGAAAACCAAGGGCGGCCATGATTGAAACAAGATGTCACGAGGTTCCTC 420
QY 421 CGGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGCAATCGGCTGCT 480
Db 421 CGGCCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGCAATCGGCTGCT 480
QY 481 CTGATGCGCGCTGTTCGGCTGTTCAGCGCAGGGGCGCCGGTCTCTTTTGTCAAGACCG 540
Db 481 CTGATGCGCGCTGTTCGGCTGTTCAGCGCAGGGGCGCCGGTCTCTTTTGTCAAGACCG 540
QY 541 ACCTGTCCGGTGCCTGAATGAACCTGACGAGGAGGAGCGGCGCTATCGTGGCTGGCCA 600
Db 541 ACCTGTCCGGTGCCTGAATGAACCTGACGAGGAGGAGCGGCGCTATCGTGGCTGGCCA 600
QY 601 CGAGCGGCTTCCCTTGGCGAGCTGTGCTCGAGCTTGTCACTGAAGCGGAGGAGCTGGC 660
Db 601 CGAGCGGCTTCCCTTGGCGAGCTGTGCTCGAGCTTGTCACTGAAGCGGAGGAGCTGGC 660
QY 661 TGCTATTGGGCAAGTGCCTGGGCGAGGATCTCTGTCTCATCTCACCTTGTCTCGCGAGA 720
Db 661 TGCTATTGGGCAAGTGCCTGGGCGAGGATCTCTGTCTCATCTCACCTTGTCTCGCGAGA 720
QY 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCATACGCTTGTGATCGGCTACCTGCC 780
Db 721 AAGTATCCATCATGCTGATGCAATGCGGCGGCTGCATACGCTTGTGATCGGCTACCTGCC 780
QY 781 CATTTCGACCAACCAAGCAACATCGCATCGAGCGAGCAGTACTCGGATGGAACCGGTC 840
Db 781 CATTTCGACCAACCAAGCAACATCGCATCGAGCGAGCAGTACTCGGATGGAACCGGTC 840
QY 841 TTGTCTGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCG 900
Db 841 TTGTCTGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCG 900
QY 901 CCAGGCTCAAGGCGCGCATCCCGAGGAGATCTGCTGTCGACCATGGCGATGCTCT 960
Db 901 CCAGGCTCAAGGCGCGCATCCCGAGGAGATCTGCTGTCGACCATGGCGATGCTCT 960
QY 961 GCTTTCGGAATATCATGCTGGAATAATGGCCGCTTTCTGATTCATCATGCTGGCGGC 1020
Db 961 GCTTTCGGAATATCATGCTGGAATAATGGCCGCTTTCTGATTCATCATGCTGGCGGC 1020
QY 1021 TGGGTGCGGGAACCGCTATCAGGATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGC 1080
Db 1021 TGGGTGCGGGAACCGCTATCAGGATAGCGTTGGCTACCCGCTGATATTGCTGAAGAGC 1080
QY 1081 TTGGCGGGAATGGGCTGACCGCTTCCCTGCTTTACCGGTATCGCGCTCCCGATTGCG 1140
Db 1081 TTGGCGGGAATGGGCTGACCGCTTCCCTGCTTTACCGGTATCGCGCTCCCGATTGCG 1140
QY 1141 AGCGCATCCGCTTCTATCCGCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACG 1200
Db 1141 AGCGCATCCGCTTCTATCCGCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACG 1200
QY 1201 GTTTCCTCTAGGGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTAACCTTACTGGC 1260
Db 1201 GTTTCCTCTAGGGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTAACCTTACTGGC 1260
QY 1261 CGAAGCGGCTTGAATAAGGCGCGGTGTGCTTGTCTATATGTTATTTTCCACCATATTG 1320
Db 1261 CGAAGCGGCTTGAATAAGGCGCGGTGTGCTTGTCTATATGTTATTTTCCACCATATTG 1320
QY 1321 CCGTCTTTTGGCAATGTGAGGCGCCGGAACCTGGCCCTGTCTTCTTGACGAGCATTCCT 1380
Db 1321 CCGTCTTTTGGCAATGTGAGGCGCCGGAACCTGGCCCTGTCTTCTTGACGAGCATTCCT 1380
QY 1381 AGGGTCTTTTCCCTCTCGCAAAAGGAATGCAAGGTCTGTTGAAATGCTGGAAGCA 1440
```





QY 3601 TATAGCTGGGAGCCGTTCAAACAGAGGTTACTACACACACCCCATACCAAAATACATC 3660  
Db 3601 TATAGCTGGGAGCCGTTCAAACAGAGGTTACTACACACACCCCATACCAAAATACATC 3660  
QY 3661 ATGGCATCATGTCCGCTGACCTGAGGTCGTCAAGAGACCTGGGTGCTGGTAGCGGA 3720  
Db 3661 ATGGCATCATGTCCGCTGACCTGAGGTCGTCAAGAGGTCGTCAAGAGACCTGGGTGCTGGTAGCGGA 3720  
QY 3721 GTCCTAGCAGCTCTGGCCGCGTATTGCCCTGTGACAAACAGCAGCGGTGCTATGTTGGGAGG 3780  
Db 3721 GTCCTAGCAGCTCTGGCCGCGTATTGCCCTGTGACAAACAGCAGCGGTGCTATGTTGGGAGG 3780  
QY 3781 ATCATCTTGTCCGAAAGCCGCGCATCAATCCCGACAGGAGTCCTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCCGCGCATCAATCCCGACAGGAGTCCTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTCGCGCTCACACCTCCCTTATCATGCAACAGGGAATCAGCTGCC 3900  
Db 3841 GATGAGATGGAAGAGTCGCGCTCACACCTCCCTTATCATGCAACAGGGAATCAGCTGCC 3900  
QY 3901 GAAACAATCAAACAGAGGCAATCGGTTGTGCAAAACAGACCAACAGCAGCGAGGCT 3960  
Db 3901 GAAACAATCAAACAGAGGCAATCGGTTGTGCAAAACAGACCAACAGCAGCGAGGCT 3960  
QY 3961 GCTGCTCCGCTGGTGAATCAAGTGGCGGACCCCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGGTGAATCAAGTGGCGGACCCCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGATACAAATATTTAGCAGGCTTGTCCACTGCTGGCTGGCAACCC 4080  
Db 4021 TGGAAATTCATCAGCGGATACAAATATTTAGCAGGCTTGTCCACTGCTGGCTGGCAACCC 4080  
QY 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCGCGCTCACACCCACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCGCGCTCACACCCACAT 4140  
QY 4141 ACCCTCCTGTTTAAATCTCTGGGGGATGGGTGGCGCCCAACTTGTCTCCCTCAGGCT 4200  
Db 4141 ACCCTCCTGTTTAAATCTCTGGGGGATGGGTGGCGCCCAACTTGTCTCCCTCAGGCT 4200  
QY 4201 GCTTCTGCTTTGCTAGCGCGCGCATCGCTGAGCGGCTGTTGGCAGCATAGCGCTTGGG 4260  
Db 4201 GCTTCTGCTTTGCTAGCGCGCGCATCGCTGAGCGGCTGTTGGCAGCATAGCGCTTGGG 4260  
QY 4261 AAGGCTCTTGTGATATTTTGGCAGGTTATGGAGCAGGSGTGGCAGCGGCTCTGGG 4320  
Db 4261 AAGGCTCTTGTGATATTTTGGCAGGTTATGGAGCAGGSGTGGCAGCGGCTCTGGG 4320  
QY 4321 TTTAAGGTATGAGCGCGGAGATGCCCTCAACGAGGACCTGGTTAACTTCTCCCTGCT 4380  
Db 4321 TTTAAGGTATGAGCGCGGAGATGCCCTCAACGAGGACCTGGTTAACTTCTCCCTGCT 4380  
QY 4381 ATCTCTCCCTTGGCGCCCTAGTCTGGGGTCTGTGGCGAGCGATATGCGTGGCGAC 4440  
Db 4381 ATCTCTCCCTTGGCGCCCTAGTCTGGGGTCTGTGGCGAGCGATATGCGTGGCGAC 4440  
QY 4441 GTGGGCCCAAGGAGGGGCTGTGAGTGAATGAACCGGCTGATAGCGGCTCTGGG 4500  
Db 4441 GTGGGCCCAAGGAGGGGCTGTGAGTGAATGAACCGGCTGATAGCGGCTCTGGG 4500  
QY 4501 GGTAAACAGCTCTCCCGACGCACTATGTGCTGAGAGCGAGCTGCAGCAGTGTCACT 4560  
Db 4501 GGTAAACAGCTCTCCCGACGCACTATGTGCTGAGAGCGAGCTGCAGCAGTGTCACT 4560  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGTAAGAGGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGTAAGAGGCTTCAACAGTGGATCAAC 4620  
QY 4621 GAGGACTGCTCCACGCCATGCTCCGGCTGCTGGCTTGAAGATGTTGGATTGATGCG 4680  
Db 4621 GAGGACTGCTCCACGCCATGCTCCGGCTGCTGGCTTGAAGATGTTGGATTGATGCG 4680  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCCGCGATTGGCGGA 4740

Db 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAGCTCCTGCGCGATTGCGGGA 4740  
QY 4741 GTCCCTCTTCTCTCATGTCAAGTGGGTACAGGAGTCTTGGCGGGCGACGGCATCATG 4800  
Db 4741 GTCCCTCTTCTCTCATGTCAAGTGGGTACAGGAGTCTTGGCGGGCGACGGCATCATG 4800  
QY 4801 CAAACACCTCCCATGTGGAGCACAGATCACCGGACATGTGAAACCGTTCATGAGG 4860  
Db 4801 CAAACACCTCCCATGTGGAGCACAGATCACCGGACATGTGAAACCGTTCATGAGG 4860  
QY 4861 ATCTGGGGCCCTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCATTAACCGGTAC 4920  
Db 4861 ATCTGGGGCCCTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCATTAACCGGTAC 4920  
QY 4921 ACCACGGGCCCTCAGCGCCCTCCCGGCGCAATATTTCTAGGCGCTGTGGCGGCTG 4980  
Db 4921 ACCACGGGCCCTCAGCGCCCTCCCGGCGCAATATTTCTAGGCGCTGTGGCGGCTG 4980  
QY 4981 GCTGCTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACTACGTGACGGGATG 5040  
Db 4981 GCTGCTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACTACGTGACGGGATG 5040  
QY 5041 ACCACTGACAACTAAAGTCCCTGTGCAAGTTCGGGCCCGCGAATTTTTCACAGAGTG 5100  
Db 5041 ACCACTGACAACTAAAGTCCCTGTGCAAGTTCGGGCCCGCGAATTTTTCACAGAGTG 5100  
QY 5101 GATGGGTGCGGTTGACAGCTAGCTCCAGCGTGGAAACCCCTCTACGGGAGGAGTGC 5160  
Db 5101 GATGGGTGCGGTTGACAGCTAGCTCCAGCGTGGAAACCCCTCTACGGGAGGAGTGC 5160  
QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATCGAGCCCGAA 5220  
Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATCGAGCCCGAA 5220  
QY 5221 CCGGAGTAGCAGTCTCACTTCCTGCTCAACGACCCCTCCACATTTACGGCGGAGACG 5280  
Db 5221 CCGGAGTAGCAGTCTCACTTCCTGCTCAACGACCCCTCCACATTTACGGCGGAGACG 5280  
QY 5281 GCTAAGCGTAGGCTGCCAGGGGATCTCCCGCTCTTGGCGAGCTCATCAGTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGCCAGGGGATCTCCCGCTCTTGGCGAGCTCATCAGTAGCCAG 5340  
QY 5341 CTGTCTGCGCCTTCTTTGAAGGCAACATGCACTACCCGTCATGACTCCCGGAGCGTGC 5400  
Db 5341 CTGTCTGCGCCTTCTTTGAAGGCAACATGCACTACCCGTCATGACTCCCGGAGCGTGC 5400  
QY 5401 CTCATCGAGCCCAACCTCTCTGGCGGAGATCTGGCGGAGGATGGCGGGAACATCACCCGCTGGAG 5460  
Db 5401 CTCATCGAGCCCAACCTCTCTGGCGGAGATGGCGGAGGATGGCGGGAACATCACCCGCTGGAG 5460  
QY 5461 TCAGAAATAAGGTAGTAAATTTTGACTCTTTGAGCCGCTCCAGCGAGGAGGATGAG 5520  
Db 5461 TCAGAAATAAGGTAGTAAATTTTGACTCTTTGAGCCGCTCCAGCGAGGAGGATGAG 5520  
QY 5521 AGGGAAGTATCCGTTCCGCGGAGATCTCTGGAGGCTCAGGAAATTCCTCAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCGCGGAGATCTCTGGAGGCTCAGGAAATTCCTCAGCGATG 5580  
QY 5581 CCCATATGGGACGCGCCGATTTACAACTCTGTTAGAGTCTCGAAGGACCCCGAC 5640  
Db 5581 CCCATATGGGACGCGCCGATTTACAACTCTGTTAGAGTCTCGAAGGACCCCGAC 5640  
QY 5641 TAGCTCCCTCAGGTGATACGGGTGTCATTTGCGGCTGCGAGGCCCTCCCGATACA 5700  
Db 5641 TAGCTCCCTCAGGTGATACGGGTGTCATTTGCGGCTGCGAGGCCCTCCCGATACA 5700  
QY 5701 CCTCCAGGAGGAGGACGCTTGTCTGTGAGAACTCTACCGTGTCTTCTGCTTGGG 5760  
Db 5701 CCTCCAGGAGGAGGACGCTTGTCTGTGAGAACTCTACCGTGTCTTCTGCTTGGG 5760  
QY 5761 GAGCTCGCCACAAAGACCTTCGGGAGTCCGAATCTGTGGCGCTCGACAGCGGACCGCA 5820





Db	1021	TGGGTGTGGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGC	1080
Qy	1081	TTGGCGGGAATGGCTGACCGCTTCTCTGCTGCTTTACGGTATCGCGCTCCCGATTGCG	1140
Db	1081	TTGGCGGGAATGGCTGACCGCTTCTCTGCTGCTTTACGGTATCGCGCTCCCGATTGCG	1140
Qy	1141	AGCGCATCGCTTCTATCGCTTCTTGACGAGTTCCTTGAGTTTAAACAGACCAACG	1200
Db	1141	AGCGCATCGCTTCTATCGCTTCTTGACGAGTTCCTTGAGTTTAAACAGACCAACG	1200
Qy	1201	GTTCCTCTAGCGGATCAATTCGCGCTCTCTGCTCCCGCTCCCGCTAAAGCTTACTGCG	1260
Db	1201	GTTCCTCTAGCGGATCAATTCGCGCTCTCTGCTCCCGCTCCCGCTAAAGCTTACTGCG	1260
Qy	1261	CGAAGCGCTTGAATAGCGCGGTGTGCTTGTCTATATATTTTCCACCATATG	1320
Db	1261	CGAAGCGCTTGAATAGCGCGGTGTGCTTGTCTATATATTTTCCACCATATG	1320
Qy	1321	CGGTCTTTTGGCAATGTAGGCGCGGAAACCTGGCCCTGCTTTTGACGAGCATTCCT	1380
Db	1321	CGGTCTTTTGGCAATGTAGGCGCGGAAACCTGGCCCTGCTTTTGACGAGCATTCCT	1380
Qy	1381	AGGGTCTTTCCGCTCTCGCAAGGAATGCAAGTCTGTGAATGCTGTGAAGGAAGCA	1440
Db	1381	AGGGTCTTTCCGCTCTCGCAAGGAATGCAAGTCTGTGAATGCTGTGAAGGAAGCA	1440
Qy	1441	GTTCCTCTGAAAGCTTCTTGAAGCAAAACCTGTAGCGACCTTTTCAGGCGAGCG	1500
Db	1441	GTTCCTCTGAAAGCTTCTTGAAGCAAAACCTGTAGCGACCTTTTCAGGCGAGCG	1500
Qy	1501	AAACCCCGCTTGGGCGACAGTGCCTCTGGGCGCAAAAGCCAGCTGTATTAAGATACCT	1560
Db	1501	AAACCCCGCTTGGGCGACAGTGCCTCTGGGCGCAAAAGCCAGCTGTATTAAGATACCT	1560
Qy	1561	GAAAGCGGCAAAACCCGAGTGCCTGTGAGTTGATAGTTGGAAAGAGTCAAA	1620
Db	1561	GAAAGCGGCAAAACCCGAGTGCCTGTGAGTTGATAGTTGGAAAGAGTCAAA	1620
Qy	1621	TGGCTCTCTCAAGGTATTCAACAGGGGCTGAAGGATGCCAGAGGTACCCCATGT	1680
Db	1621	TGGCTCTCTCAAGGTATTCAACAGGGGCTGAAGGATGCCAGAGGTACCCCATGT	1680
Qy	1681	ATGGGATCTGATCTGGGCGCTCGTGACATGCTTTACATGTTTGTAGTGGAGTTAAA	1740
Db	1681	ATGGGATCTGATCTGGGCGCTCGTGACATGCTTTACATGTTTGTAGTGGAGTTAAA	1740
Qy	1741	AACGTCTAGGCGCGCGCAACCGGACGAGTGTTCCTTTGAAACACGATAATACC	1800
Db	1741	AACGTCTAGGCGCGCGCAACCGGACGAGTGTTCCTTTGAAACACGATAATACC	1800
Qy	1801	ATGGCGCTATTAGGCGCTACTCCCAACAGACGCGAGGCTACTTGGCTGATCATCT	1860
Db	1801	ATGGCGCTATTAGGCGCTACTCCCAACAGACGCGAGGCTACTTGGCTGATCATCT	1860
Qy	1861	AGCCTCAAGCGCGGACAGGAAACAGGTGAGGGGAGGTCCAGTGTCTCCACCGCA	1920
Db	1861	AGCCTCAAGCGCGGACAGGAAACAGGTGAGGGGAGGTCCAGTGTCTCCACCGCA	1920
Qy	1921	ACACAACTTTCTGGGCGCTGCTCAATGCGGTGTGTGGACTGTCTATCATGGTGC	1980
Db	1921	ACACAACTTTCTGGGCGCTGCTCAATGCGGTGTGTGGACTGTCTATCATGGTGC	1980
Qy	1981	GGCTCAAAGACCTTTGCGGCGCCAAAGGCGCCCAATCAACCAATGTACCAATGTGAC	2040
Db	1981	GGCTCAAAGACCTTTGCGGCGCCAAAGGCGCCCAATCAACCAATGTGAC	2040
Qy	2041	CAGGACCTGTGCGGTGGCAAGCGCCCGCGGCGGCTTCTTGACACCATGACCTGC	2100
Db	2041	CAGGACCTGTGCGGTGGCAAGCGCCCGCGGCGGCTTCTTGACACCATGACCTGC	2100
Qy	2101	GGCAGCTCGGACCTTTACTTGTGTCAGAGGCAATGCGATGTCAATTCGGTGGCGCGG	2160
Db	2101	GGCAGCTCGGACCTTTACTTGTGTCAGAGGCAATGCGATGTCAATTCGGTGGCGCGG	2160
Qy	2161	GGCGACAGCGGGAGCCTACTCTCCCGCAGGCGCGCTCTCTACTTGAAGGCTCTTCG	2220
Db	2161	GGCGACAGCGGGAGCCTACTCTCCCGCAGGCGCGCTCTCTACTTGAAGGCTCTTCG	2220
Qy	2221	GGCGGTCCATGCTCTGCGCTCGGGCAGCTGTGGGCACTTTTCGGGCTGCGGTGTC	2280
Db	2221	GGCGGTCCATGCTCTGCGCTCGGGCAGCTGTGGGCACTTTTCGGGCTGCGGTGTC	2280
Qy	2281	ACCCGAGGGTTCGAAGCGGTGACCTTTGTACCCGTGAGTCTATGGAACCACTATG	2340
Db	2281	ACCCGAGGGTTCGAAGCGGTGACCTTTGTACCCGTGAGTCTATGGAACCACTATG	2340
Qy	2341	CGGTCCCGGCTTTCACGGACAACTCGTCCCTCGCGCGTACCCGACGATTCAGGTG	2400
Db	2341	CGGTCCCGGCTTTCACGGACAACTCGTCCCTCGCGCGTACCCGACGATTCAGGTG	2400
Qy	2401	GGCCATCTACAGCGCCCTACTGGTAGCGCAAGACCTAAGGTGCGGCTGGTATGCA	2460
Db	2401	GGCCATCTACAGCGCCCTACTGGTAGCGCAAGACCTAAGGTGCGGCTGGTATGCA	2460
Qy	2461	GGCCAGGGTATAAGGTGCTTCTGTAACCCGTCCGTGCGCGCACTAGGTTTCGGG	2520
Db	2461	GGCCAGGGTATAAGGTGCTTCTGTAACCCGTCCGTGCGCGCACTAGGTTTCGGG	2520
Qy	2521	GGGTATATCTTAAGGACATGGTATCGACCTTAACATCAGAACCGGGTAAAGCAATC	2580
Db	2521	GGGTATATCTTAAGGACATGGTATCGACCTTAACATCAGAACCGGGTAAAGCAATC	2580
Qy	2581	ACCACGGGTCCCGCATACAGTACTCCACTTGGCAAGTTTCTGCGGACGGTGGTTC	2640
Db	2581	ACCACGGGTCCCGCATACAGTACTCCACTTGGCAAGTTTCTGCGGACGGTGGTTC	2640
Qy	2641	TCCTGGGCGGCTTACATCATATAATGTGATGAGTGCCACTCAACTGACTCGACCAT	2700
Db	2641	TCCTGGGCGGCTTACATCATATAATGTGATGAGTGCCACTCAACTGACTCGACCAT	2700
Qy	2701	ATCTGGGCACTCGGCAAGTCCCTGGACCAAGCGGAGCGGCTGGAGCGGCTGCTGTG	2760
Db	2701	ATCTGGGCACTCGGCAAGTCCCTGGACCAAGCGGAGCGGCTGGAGCGGCTGCTGTG	2760
Qy	2761	CTGCGCACCGCTTACGCTCCGGATCGGTCAACCGTGCACATCCAAACATCGAGAGGTG	2820
Db	2761	CTGCGCACCGCTTACGCTCCGGATCGGTCAACCGTGCACATCCAAACATCGAGAGGTG	2820
Qy	2821	GCTCTGTCAGCACTGAGAAATCCCTTTTATGGCAAGCAATCCCATCGAGACCATC	2880
Db	2821	GCTCTGTCAGCACTGAGAAATCCCTTTTATGGCAAGCAATCCCATCGAGACCATC	2880
Qy	2881	AAGGGGGGAGGCACTCATTTCTGCAATTCGAAGAAATGTGATGAGTCCGCGG	2940
Db	2881	AAGGGGGGAGGCACTCATTTCTGCAATTCGAAGAAATGTGATGAGTCCGCGG	2940
Qy	2941	AAGCTGTCCGCGCTCGGACTCAATGCTGAGCAATTAACCGGGGCTTGTATGATCCGTC	3000
Db	2941	AAGCTGTCCGCGCTCGGACTCAATGCTGAGCAATTAACCGGGGCTTGTATGATCCGTC	3000
Qy	3001	ATACCACTAGCGGAGAGTCAATTTCTGCAATTCGAAGAAATGTGATGAGTCCGCGG	3060
Db	3001	ATACCACTAGCGGAGAGTCAATTTCTGCAATTCGAAGAAATGTGATGAGTCCGCGG	3060
Qy	3061	GGCGATTTTCGACTCAGTGTGATGCAATCATGTGTCAACCGAGAGTGCACCTCAGC	3120
Db	3061	GGCGATTTTCGACTCAGTGTGATGCAATCATGTGTCAACCGAGAGTGCACCTCAGC	3120
Qy	3121	CTGGAACCGGCTTACCATTTGAGACGAGCCGTGCGCAAGACCGGCTGTCAGCTCG	3180
Db	3121	CTGGAACCGGCTTACCATTTGAGACGAGCCGTGCGCAAGACCGGCTGTCAGCTCG	3180
Qy	3181	CAGCGCGGAGGAGGAGTGGTGGGCGAGATGGGCAATTAAGTTTGTGACTCCAGGA	3240
Db	3181	CAGCGCGGAGGAGGAGTGGTGGGCGAGATGGGCAATTAAGTTTGTGACTCCAGGA	3240

QY 3241 GAACGGCCCTCGGGGATGTTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGCTGT 3300  
Db 3241 GAACGGCCCTCGGGGATGTTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGCTGT 3300  
QY 3301 GCTTGTACGAGCTACGCGCCGCGAGACCTCAGTTAGTTGCGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTACGAGCTACGCGCCGCGAGACCTCAGTTAGTTGCGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTTCCCGTCTGCGCAGGACCAATCTGGAGTTCTGGGAGAGCGTCTTTACAGCCCTC 3420  
Db 3361 CCAGGTTTCCCGTCTGCGCAGGACCAATCTGGAGTTCTGGGAGAGCGTCTTTACAGCCCTC 3420  
QY 3421 ACCCATAGACGCCCATTTCTTTGTCACAGACTAAGCAGGACGAGCAAACTTCCCTAC 3480  
Db 3421 ACCCATAGACGCCCATTTCTTTGTCACAGACTAAGCAGGACGAGCAAACTTCCCTAC 3480  
QY 3481 CTGGTAGCATACGAGCTACGCTGAGGTTGCGCCAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
Db 3481 CTGGTAGCATACGAGCTACGCTGAGGTTGCGCCAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACTACACACACCCCAACCAATACATC 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACTACACACACCCCAACCAATACATC 3600  
QY 3601 TATAGGCTGGAGCGGTTTCAAAACGAGTTTACTACACACACCCCAACCAATACATC 3660  
Db 3601 TATAGGCTGGAGCGGTTTCAAAACGAGTTTACTACACACACCCCAACCAATACATC 3660  
QY 3661 ATGCAATCATGCTGGCTGAGCTGAGTGTCTGAGGACACTGGGTCTGTTAGGCGGA 3720  
Db 3661 ATGCAATCATGCTGGCTGAGCTGAGTGTCTGAGGACACTGGGTCTGTTAGGCGGA 3720  
QY 3721 GTCCTAGCAGCTCTGGCGCGGTATGCTGACAAACAGCAGCGGTGTCATGTTGGCAGG 3780  
Db 3721 GTCCTAGCAGCTCTGGCGCGGTATGCTGACAAACAGCAGCGGTGTCATGTTGGCAGG 3780  
QY 3781 ATCATTTGTCGGAAACCGCGCATCATTTCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
Db 3781 ATCATTTGTCGGAAACCGCGCATCATTTCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGAAGTGCCTCACAACCTGCTTACATCGAAACAGGGAATGCACTCGCC 3900  
Db 3841 GATGAGATGAAGTGCCTCACAACCTGCTTACATCGAAACAGGGAATGCACTCGCC 3900  
QY 3901 GAAACAAATTCACACAGAGGCAATCGGTTCTGCAACACGCCACCAAGACGAGGCT 3960  
Db 3901 GAAACAAATTCACACAGAGGCAATCGGTTCTGCAACACGCCACCAAGACGAGGCT 3960  
QY 3961 GCTGCTCCGCTGTGGATTCACAGTGGCGGACCTCGAAGCCTTCTGGGCAAGCATATG 4020  
Db 3961 GCTGCTCCGCTGTGGATTCACAGTGGCGGACCTCGAAGCCTTCTGGGCAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCACTGTGCTGCAACCCC 4080  
Db 4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCACTGTGCTGCAACCCC 4080  
QY 4081 GCGATAGCATCACTGATGCAATTCACAGCTCTATCACCAGCCGCTCACCACCCCAAT 4140  
Db 4081 GCGATAGCATCACTGATGCAATTCACAGCTCTATCACCAGCCGCTCACCACCCCAAT 4140  
QY 4141 ACCCTCTCTTTAAACATCTGGGGGATGGTGGCGGCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCTCTTTAAACATCTGGGGGATGGTGGCGGCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTGCTTTCTAGGCGCGGATCGTGTAGAGCGGTGTTGGAGCATAGGCTTGGG 4260  
Db 4201 GCTTCTGCTTTCTAGGCGCGGATCGTGTAGAGCGGTGTTGGAGCATAGGCTTGGG 4260  
QY 4261 AAGGTGCTTGTGGATTTTGGCAGTTATGAGCAGGGTGGCAGGCGCTCGTGGCC 4320  
Db 4261 AAGGTGCTTGTGGATTTTGGCAGTTATGAGCAGGGTGGCAGGCGCTCGTGGCC 4320  
QY 4321 TTTAAGGTGATGAGCGGCGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380

Db 4321 TTTAAGGTGATGAGCGGCGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGGCGCCCTAGTCTGTCGGGTCGTGTGCGAGCGATCTGCGTTCGGAC 4440  
Db 4381 ATCTCTCTCCCTGGCGCCCTAGTCTGTCGGGTCGTGTGCGAGCGATCTGCGTTCGGAC 4440  
QY 4441 GTGGGCCACGAGGAGGGGCTGTGCAGTGGATGAAACCGGCTGATAGCGTTCGTTCCGG 4500  
Db 4441 GTGGGCCACGAGGAGGGGCTGTGCAGTGGATGAAACCGGCTGATAGCGTTCGTTCCGG 4500  
QY 4501 GGTAAACAGGCTCTCCCCACGCACTATGTGCTGAGAGCGAGCTGACAGCGTGTCACT 4560  
Db 4501 GGTAAACAGGCTCTCCCCACGCACTATGTGCTGAGAGCGAGCTGACAGCGTGTCACT 4560  
QY 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGCTGTCTGAAAGAGGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGCTGTCTGAAAGAGGCTTCAACAGTGGATCAAC 4620  
QY 4621 GAGGACTGCTCCAGCCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
Db 4621 GAGGACTGCTCCAGCCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
QY 4681 ACAGTGTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAAAGCTCTGCGCGGATTCGCGG 4740  
Db 4681 ACAGTGTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAAAGCTCTGCGCGGATTCGCGG 4740  
QY 4741 GTCCCTCTCTCTCATGTCACAGTGGGTACAGGAGTCTGCGCGGCGGAGCGGATCATG 4800  
Db 4741 GTCCCTCTCTCTCATGTCACAGTGGGTACAGGAGTCTGCGCGGCGGAGCGGATCATG 4800  
QY 4801 CAAACACCTGCTCCCATGTTGAGGACACAGATCACCGGACATGTGAAACAGGTTCCATGAGG 4860  
Db 4801 CAAACACCTGCTCCCATGTTGAGGACACAGATCACCGGACATGTGAAACAGGTTCCATGAGG 4860  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCATTTAACCGGTAC 4920  
Db 4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCATTTAACCGGTAC 4920  
QY 4921 ACCACGGGCTTGCACGCTTCCCGGCGGCGGCGGCGGCGGCGGCGGCTGTGCGGCGGT 4980  
Db 4921 ACCACGGGCTTGCACGCTTCCCGGCGGCGGCGGCGGCGGCGGCGGCTGTGCGGCGGT 4980  
QY 4981 GCTGCTGAGGAGTACGTGGAGTTACGCGGTCGGGATTTCCACACTGTCAGCGGCGATG 5040  
Db 4981 GCTGCTGAGGAGTACGTGGAGTTACGCGGTCGGGATTTCCACACTGTCAGCGGCGATG 5040  
QY 5041 ACCACTGACAACTGAAAGTCCGCTGTCAGGTTCCGGGCGGCGGCGGCGGCGGCGGCGG 5100  
Db 5041 ACCACTGACAACTGAAAGTCCGCTGTCAGGTTCCGGGCGGCGGCGGCGGCGGCGGCGG 5100  
QY 5101 GATGGGTCGGGTGACAGGTCAGCTCCAGCGTCAAGACCCCTCTCTAGCGGAGGCTC 5160  
Db 5101 GATGGGTCGGGTGACAGGTCAGCTCCAGCGTCAAGACCCCTCTCTAGCGGAGGCTC 5160  
QY 5161 ACATTTCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGTCCCATGCGGCGGCGGCGG 5220  
Db 5161 ACATTTCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGTCCCATGCGGCGGCGGCGG 5220  
QY 5221 CCGGAGTAGAGTGTCACTTCCATGCTACCGACCCCTCCCATTTACGCGGAGAGCG 5280  
Db 5221 CCGGAGTAGAGTGTCACTTCCATGCTACCGACCCCTCCCATTTACGCGGAGAGCG 5280  
QY 5281 GCTAAGCGTAGGTCGGCGAGGATCTCCCGCTCTGTCGCGAGCTCATAGCTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGTCGGCGAGGATCTCCCGCTCTGTCGCGAGCTCATAGCTAGCCAG 5340  
QY 5341 CTGTCTGCGCTTCTTGAAGCAACATGCACTACCCCTCATGCTCCCGGAGCGCTGAC 5400  
Db 5341 CTGTCTGCGCTTCTTGAAGCAACATGCACTACCCCTCATGCTCCCGGAGCGCTGAC 5400  
QY 5401 CTCTAGGCGCAACCTCTCTGCGGCGAGGATGGCGGGAACATCACCGCGGTGAG 5460

5401	Db	CTCATCGAGCCAACTCTCTGTGGCGGACGAGAGATGGCGGGAACAATCACCCGCGTGGAG	5460	QY	6541	ATGGGCTCTTCAVACGGAATCCAAATACTCTCTCTGGACAGCGGGTTCGAGTTCCTGTTGAAT	6600
5461	QY	TCAGAAATAGGTAGTAATTTTGGACTCTTTTCGAGCGCTTCCAAAGCGGAGAGATGAG	5520	Db	6541	ATGGGCTCTTCAVACGGAATCCAAATACTCTCTCTGGACAGCGGGTTCGAGTTCCTGTTGAAT	6600
5461	Db	TCAGAAATAGGTAGTAATTTTGGACTCTTTTCGAGCGCTTCCAAAGCGGAGAGATGAG	5520	QY	6601	GCCTGGAAGCGAAGAAATGCCCTATGGGCTTCGCATATGACACCCGCTGTTTGTACTCA	6660
5521	QY	AGGGAAGTATCCGTTCCGGCGGAGATCTCTCGGAGGTCCAGGAATTCCTTCGAGCGATG	5580	Db	6601	GCCTGGAAGCGAAGAAATGCCCTATGGGCTTCGCATATGACACCCGCTGTTTGTACTCA	6660
5521	Db	AGGGAAGTATCCGTTCCGGCGGAGATCTCTCGGAGGTCCAGGAATTCCTTCGAGCGATG	5580	QY	6661	ACGGTCACTGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATGTTGTGACTTGCGC	6720
5581	QY	CCCATATGGGCACGCGCCGGATTAACCCCTCCACTGTTAGAGTCTCGAAGACCCGGAC	5640	Db	6661	ACGGTCACTGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATGTTGTGACTTGCGC	6720
5581	Db	CCCATATGGGCACGCGCCGGATTAACCCCTCCACTGTTAGAGTCTCGAAGACCCGGAC	5640	QY	6721	CCCGAAGCCACAGAGGCGCATTAAGGTGCTCACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780
5641	QY	TAGTCTCCCTCAGTGTACAGGGGTGTCATTTGCCGCTGCGAAGGCCCTTCGGATACCA	5700	Db	6721	CCCGAAGCCACAGAGGCGCATTAAGGTGCTCACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780
5641	Db	TAGTCTCCCTCAGTGTACAGGGGTGTCATTTGCCGCTGCGAAGGCCCTTCGGATACCA	5700	QY	6781	ACTAATTTCTAAAGGGCAGAACTGGCGTATCGCGGTGCGCGGAGCGGTGTACTGACG	6840
5701	QY	CCTCCAGGAGGAGGACGGTTCCTGTGAGAAATCTACCGTGTCTTCGCTTGGCG	5760	Db	6781	ACTAATTTCTAAAGGGCAGAACTGGCGTATCGCGGTGCGCGGAGCGGTGTACTGACG	6840
5701	Db	CCTCCAGGAGGAGGACGGTTCCTGTGAGAAATCTACCGTGTCTTCGCTTGGCG	5760	QY	6841	ACCAGTCCGCGTAATACCCCTCACATGTTACTTTGAAGCCGCTGCGGCTGTGAGCTGG	6900
5761	QY	GAGTCTGCCACAAAGACCTTCGGCAGCTCCGAATCGTGGCGCTGCGACGCGCACGGCA	5820	Db	6841	ACCAGTCCGCGTAATACCCCTCACATGTTACTTTGAAGCCGCTGCGGCTGTGAGCTGG	6900
5761	Db	GAGTCTGCCACAAAGACCTTCGGCAGCTCCGAATCGTGGCGCTGCGACGCGCACGGCA	5820	QY	6901	AAGCTCCAGGACTGCAAGATGCTGTATGCGGAGACGACCTTGTGCTTATCTGTGAAGC	6960
5821	QY	ACGGCTCTCTGACAGGCGCTCCGACGCGGAGTCCGAGTTCGAGTTCGTATC	5880	Db	6901	AAGCTCCAGGACTGCAAGATGCTGTATGCGGAGACGACCTTGTGCTTATCTGTGAAGC	6960
5821	Db	ACGGCTCTCTGACAGGCGCTCCGACGCGGAGTCCGAGTTCGAGTTCGTATC	5880	QY	6961	GGGGGACCCAAAGAGGACGAGCGGAGCTTACGGGCTTCCAGGAGGCTAGTAGTATAC	7020
5881	QY	TCTTCCATGCCCCCTTGGAGGGGAGCGGGGGATCCCGATCTCAGCGACGGGCTTGG	5940	Db	6961	GGGGGACCCAAAGAGGACGAGCGGAGCTTACGGGCTTCCAGGAGGCTAGTAGTATAC	7020
5881	Db	TCTTCCATGCCCCCTTGGAGGGGAGCGGGGGATCCCGATCTCAGCGACGGGCTTGG	5940	QY	7021	TCGTGCCCCCTTGGGAGCCGCCCCAAACCCAGAAATACGACTTGGAGTGTAAATCATG	7080
5941	QY	TCTACCGTAAAGCGAGAGGCTAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6000	Db	7021	TCGTGCCCCCTTGGGAGCCGCCCCAAACCCAGAAATACGACTTGGAGTGTAAATCATG	7080
5941	Db	TCTACCGTAAAGCGAGAGGCTAGTGAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6000	QY	7081	TCCTCCAAATGTGTAGTTCGCGACCGATGATCTGGGAAAAGGGTGTACTTCTCACCCGT	7140
6001	QY	ACAGGCGCCCTGATACGCGCATGGCTGCGGAGGAAACCAAGTGCCTCATGACTG	6060	Db	7081	TCCTCCAAATGTGTAGTTCGCGACCGATGATCTGGGAAAAGGGTGTACTTCTCACCCGT	7140
6001	Db	ACAGGCGCCCTGATACGCGCATGGCTGCGGAGGAAACCAAGTGCCTCATGACTG	6060	QY	7141	GACCCACCAACCCCTTGGCGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT	7200
6061	QY	AGCAACTCTTTGCTCGCTCACCAAACTTGGTCTATGCTGCTGCTGCTGCTGCTGCTG	6120	Db	7141	GACCCACCAACCCCTTGGCGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT	7200
6061	Db	AGCAACTCTTTGCTCGCTCACCAAACTTGGTCTATGCTGCTGCTGCTGCTGCTGCTG	6120	QY	7201	TCCTGCTAGGCAACATCATGATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
6121	QY	CTGCGGAGAGAGAGTCACTTTGACAGACTGCGGTCTGCGAGGAAACCAAGTGCCTCATG	6180	Db	7201	TCCTGCTAGGCAACATCATGATGCGCCACCTTGTGGCAAGGATGATCCTGATG	7260
6121	Db	CTGCGGAGAGAGAGTCACTTTGACAGACTGCGGTCTGCGAGGAAACCAAGTGCCTCATG	6180	QY	7261	ACTCATTTCTTCTCCATCTTCTAGCTCAGGAAACAACTTGGAAAAGCCCTAGATTGTG	7320
6181	QY	GTGCTCAAGGAGATGAAGGCGAAGCGTCCACAGTTAAGGCTAACTTCTATCCGTGGAG	6240	Db	7261	ACTCATTTCTTCTCCATCTTCTAGCTCAGGAAACAACTTGGAAAAGCCCTAGATTGTG	7320
6181	Db	GTGCTCAAGGAGATGAAGGCGAAGCGTCCACAGTTAAGGCTAACTTCTATCCGTGGAG	6240	QY	7321	ATCTACGGGCTTGTACTTCTCATTGAGCCACTTGACCTACTCAGATCAATCAACGACTC	7380
6241	QY	GAGGCTGTAAAGTGAAGGCGGCGGATTCGCGGAGTCTAAATTTGGCTATGGGCAAG	6300	Db	7321	ATCTACGGGCTTGTACTTCTCATTGAGCCACTTGACCTACTCAGATCAATCAACGACTC	7380
6241	Db	GAGGCTGTAAAGTGAAGGCGGCGGATTCGCGGAGTCTAAATTTGGCTATGGGCAAG	6300	QY	7381	CATGGCTTAGCGCATTTTCACTCCATAGTTACTTCTCAGGTGAGATCAATAGGGTGGCT	7440
6301	QY	GAGTCCGGAACTTATCCAGCAAGGCGTAAACACATCCGCTCCGTTGTGGAAGACTTG	6360	Db	7381	CATGGCTTAGCGCATTTTCACTCCATAGTTACTTCTCAGGTGAGATCAATAGGGTGGCT	7440
6301	Db	GAGTCCGGAACTTATCCAGCAAGGCGTAAACACATCCGCTCCGTTGTGGAAGACTTG	6360	QY	7441	TCATGCTCAGGAAAATTTGGGGTACCGGCCCTTTCGAGTCTGGAGACATCGGGCCAGAGT	7500
6361	QY	CTGGAAGACACTGAGACCAATTTGACACCAATCATGCGCAAAATAGGTTTCTG	6420	Db	7441	TCATGCTCAGGAAAATTTGGGGTACCGGCCCTTTCGAGTCTGGAGACATCGGGCCAGAGT	7500
6361	Db	CTGGAAGACACTGAGACCAATTTGACACCAATCATGCGCAAAATAGGTTTCTG	6420	QY	7501	GTCCGCGCTAGGCTTACTGTCCAGGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
6421	QY	GTCCAAACAGAGAGGCGGCGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6480	Db	7501	GTCCGCGCTAGGCTTACTGTCCAGGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC	7560
6421	Db	GTCCAAACAGAGAGGCGGCGGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6480	QY	7561	AACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCTCCAGTGGAT	7620
6481	QY	GTTCTGTGTGCGAAGAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCCGTG	6540	Db	7561	AACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCTCCAGTGGAT	7620
6481	Db	GTTCTGTGTGCGAAGAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCCGTG	6540				



QY 7621 TTATCAGCTGGTTCCTGCTGTTACAGCGGGGAGACATATATACAGCCCTGTCTCCT 7680  
Db |||||  
QY 7621 TTATCAGCTGGTTCCTGCTGTTACAGCGGGGAGACATATATACAGCCCTGTCTCCT 7680  
Db |||||  
QY 7681 GCCGACCCCGCTGGTTCATGTCCTTACTCTACTCTTCTGTAGGGGTAGCACTAT 7740  
Db |||||  
QY 7681 GCCGACCCCGCTGGTTCATGTCCTTACTCTACTCTTCTGTAGGGGTAGCACTAT 7740  
Db |||||  
QY 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800  
Db |||||  
QY 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800  
Db |||||  
QY 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7921 TAGCTGTGAAGGTCCTGTAGCGGCTTGACTGCAGAGAGTCTGATCTGGCTCTCTGC 7980  
Db |||||  
QY 7921 TAGCTGTGAAGGTCCTGTAGCGGCTTGACTGCAGAGAGTCTGATCTGGCTCTCTGC 7980  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||  
RESULT 10  
ADP86271  
ID ADP86271 standard; DNA; 11313 BP.  
XX  
AC ADP86271;  
XX  
DT 23-SEP-2004 (first entry)  
XX  
DE Hepatitis C virus Con-1 replicon I377/NS3-3' derived plasmid DNA, pZSL0.  
XX  
KW Hepatitis C virus; HCV; anti-HCV agent; HCV infection; therapy; plasmid;  
KW ds.  
XX  
OS Hepatitis C virus.  
XX  
FN W0200405216-A2.  
XX  
PD 01-JUL-2004.  
XX  
PF 12-DEC-2003; 2003WO-US039722.  
XX  
PR 13-DEC-2002; 2002US-0433303P.  
XX  
PA (FOXC-) FOX CHASE CANCER CENT.  
XX  
PI Zhu Q, Guo J, Seeger C;  
XX  
WPI; 2004-488079/46.  
XX  
New cell-line that replicates hepatitis C virus (HCV), where the cell  
line is selected from a non-human cell line and a human non-hepatic cell  
line, useful for identifying anti-HCV agents for treating HCV infections.  
Example III; SEQ ID NO 8; 130pp; English.  
XX  
The present invention provides hepatitis C virus (HCV) replication cells  
and cell lines derived from human non-hepatic cells or non-human cells.  
The invention is useful for identifying anti-HCV agents for treating HCV  
infections. The present sequence is hepatitis C virus Con-1 replicon  
I377/NS3-3' derived plasmid DNA.  
XX  
SQ Sequence 11313 BP; 2346 A; 3334 C; 3153 G; 2480 T; 0 U; 0 Other;  
Query Match 100.0%; Score 7987.4; DB 12; Length 11313;

Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCAGCCCCCGATTGGGGGACACTCCACCATAGATCATCTCCCTGTGAGGAACACTG 60  
Db |||||  
QY 1 GCAGCCCCCGATTGGGGGACACTCCACCATAGATCATCTCCCTGTGAGGAACACTG 60  
Db |||||  
QY 61 TCTTCAAGCAGAAAGCGTCTAGCCATGGGTTAGTATAGTGTGCTGAGAGCTTCCAGGAC 120  
Db |||||  
QY 61 TCTTCAAGCAGAAAGCGTCTAGCCATGGGTTAGTATAGTGTGCTGAGAGCTTCCAGGAC 120  
Db |||||  
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGAG 180  
Db |||||  
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGAG 180  
Db |||||  
QY 181 GACGACCGGGTCTTTCTTGGATCAACCGGTCAATGCTTGAGATTTGGGGTGGCCCC 240  
Db |||||  
QY 181 GACGACCGGGTCTTTCTTGGATCAACCGGTCAATGCTTGAGATTTGGGGTGGCCCC 240  
Db |||||  
QY 241 GCGAGACTGCTAGCCAGTAGTGTGGGTGCGAAAGCGCTTGTGCTATCTGCTCATAGG 300  
Db |||||  
QY 241 GCGAGACTGCTAGCCAGTAGTGTGGGTGCGAAAGCGCTTGTGCTATCTGCTCATAGG 300  
Db |||||  
QY 301 GTGCTTGGAGTGTCCCGGAGGCTCTGTAGACCGTGCACCATGAGCAGATCCTAAAC 360  
Db |||||  
QY 301 GTGCTTGGAGTGTCCCGGAGGCTCTGTAGACCGTGCACCATGAGCAGATCCTAAAC 360  
Db |||||  
QY 361 CTAAAGAAAACCAAGGGCGCCCATGATTGAACAGATGGATTGACGAGGTTCTC 420  
Db |||||  
QY 361 CTAAAGAAAACCAAGGGCGCCCATGATTGAACAGATGGATTGACGAGGTTCTC 420  
Db |||||  
QY 421 CGCCCGCTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAGACATTCGGCTGCT 480  
Db |||||  
QY 421 CGCCCGCTGGGTGAGAGGCTATTGCGCTATGACTGGGCAACAGACATTCGGCTGCT 480  
Db |||||  
QY 481 CTGATCGCGGCTGTTCCGGCTGTCAGCGAGGGCGCCGGTCTTTTGTCAAGACCG 540  
Db |||||  
QY 481 CTGATCGCGGCTGTTCCGGCTGTCAGCGAGGGCGCCGGTCTTTTGTCAAGACCG 540  
Db |||||  
QY 541 ACCTGTCGGTGCCTGAATGAATCTGAGCAGGAGCGCGGCTATCGTGGTGGCCA 600  
Db |||||  
QY 541 ACCTGTCGGTGCCTGAATGAATCTGAGCAGGAGCGCGGCTATCGTGGTGGCCA 600  
Db |||||  
QY 601 CGACGGCGTTCCTTGGCAGCTGTGTCACAGTGTGTCAGTGAAGCGGAGGAGCTGGC 660  
Db |||||  
QY 601 CGACGGCGTTCCTTGGCAGCTGTGTCACAGTGTGTCAGTGAAGCGGAGGAGCTGGC 660  
Db |||||  
QY 661 TGCTATTGGCGAAGTCCCGGGCAGGATCTCTGTCTATCTACCTTGTCTTCCGCCAGA 720  
Db |||||  
QY 661 TGCTATTGGCGAAGTCCCGGGCAGGATCTCTGTCTATCTACCTTGTCTTCCGCCAGA 720  
Db |||||  
QY 721 AAGTATCCATCATGGCTGATGCAATGGGGGCTGATACGCTTGTATCGGCTACCTGCC 780  
Db |||||  
QY 721 AAGTATCCATCATGGCTGATGCAATGGGGGCTGATACGCTTGTATCGGCTACCTGCC 780  
Db |||||  
QY 781 CATTCGACCACCAAGCGAAACATCGCATCGAGCAGCAGTCTCGATCGAGACCGGTC 840  
Db |||||  
QY 781 CATTCGACCACCAAGCGAAACATCGCATCGAGCAGCAGTCTCGATCGAGACCGGTC 840  
Db |||||  
QY 841 TTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCGAAGCTTTCG 900  
Db |||||  
QY 841 TTGTCGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCAGCGAAGCTTTCG 900  
Db |||||  
QY 901 CCAGGCTCAAGCGCGCATGCCCGAGGAGGATCTCGTCTGACCATGCGATGCTC 960  
Db |||||  
QY 901 CCAGGCTCAAGCGCGCATGCCCGAGGAGGATCTCGTCTGACCATGCGATGCTC 960  
Db |||||  
QY 961 GCTTCCGGAATATCATGTGGAAATGCGCGTTTCTTGGATTCTACACTGTGCGCGC 1020  
Db |||||  
QY 961 GCTTCCGGAATATCATGTGGAAATGCGCGTTTCTTGGATTCTACACTGTGCGCGC 1020  
Db |||||  
QY 1021 TGGGTGTGCGGACCGCTTATCAGGACATAGCGTTGGCTTACCGGTGATATTGCTGAGAGC 1080  
Db |||||

Db	1021	TGGGTGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCGGTGATATTGCTGAAGAGC	1080	Qy	2161	GGCGACAGCGGGAGCGCTACTCTCCCCAGGCCCGTCTCTCTACTTAAGAGGGCTCTTCG	2220
Qy	1081	TTGCGCGGGAATGGGCTGACCGCTTCCTGCTGTTTACGGTATCGCGCTCCCGATTTCGC	1140	Db	2161	GGCGACAGCGGGAGCGCTACTCTCCCCAGGCCCGTCTCTCTACTTAAGAGGGCTCTTCG	2220
Db	1081	TTGCGCGGGAATGGGCTGACCGCTTCCTGCTGTTTACGGTATCGCGCTCCCGATTTCGC	1140	Qy	2221	GGCGGTCCACTGCTCTGCGGTCGCGGACGCTGTGGGCATCTTTTCGGGGCTGCGGTGTC	2280
Qy	1141	AGCGATCGCCTTCTATCGCCTTCTTGACGATTCCTTGAGTTCTTGAAGTTAAACAGCACCAACG	1200	Db	2221	GGCGGTCCACTGCTCTGCGGTCGCGGACGCTGTGGGCATCTTTTCGGGGCTGCGGTGTC	2280
Db	1141	AGCGATCGCCTTCTATCGCCTTCTTGACGATTCCTTGAGTTCTTGAAGTTAAACAGCACCAACG	1200	Qy	2281	ACCGAGGGGTGGGAAGGGCGGTGACTTTGTATCCCGTTCGAGTCTATCGAAACCACTATG	2340
Qy	1201	GTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAGTTTACTGCGC	1260	Db	2281	ACCGAGGGGTGGGAAGGGCGGTGACTTTGTATCCCGTTCGAGTCTATCGAAACCACTATG	2340
Db	1201	GTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAAGTTTACTGCGC	1260	Qy	2341	CGGTCCCGGTCTTACCGACAACCTGTCCTCGCGCGTACCGACACATTCCAGGTG	2400
Qy	1261	CGAAGCGGCTTGAATAAGGCGGTGTGCTTGTCTATATGTTATTTTCAACCAATTTG	1320	Db	2341	CGGTCCCGGTCTTACCGACAACCTGTCCTCGCGCGTACCGACACATTCCAGGTG	2400
Db	1261	CGAAGCGGCTTGAATAAGGCGGTGTGCTTGTCTATATGTTATTTTCAACCAATTTG	1320	Qy	2401	GCCCATCTACAGCGCCCTACTGTTAGCGGCAAGACACTAAGTTCGCGGTGCGTATGCA	2460
Qy	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGCTCTTGACAGGCAATTCCT	1380	Db	2401	GCCCATCTACAGCGCCCTACTGTTAGCGGCAAGACACTAAGTTCGCGGTGCGTATGCA	2460
Db	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGCTCTTGACAGGCAATTCCT	1380	Qy	2461	GCCCAAGGATATAAGGTGCTTCTCTGAAACCGTCCGTCGCGCGCACCCCTAGSTTTTCGGG	2520
Qy	1381	AGGGGTCTTTCCCTCTGCGCAAAAGGATGCAAGGTCTGTTGAATGTCGTGAAGGAACA	1440	Db	2461	GCCCAAGGATATAAGGTGCTTCTCTGAAACCGTCCGTCGCGCGCACCCCTAGSTTTTCGGG	2520
Db	1381	AGGGGTCTTTCCCTCTGCGCAAAAGGATGCAAGGTCTGTTGAATGTCGTGAAGGAACA	1440	Qy	2521	GCGTATATGTTAAAGGCAATGTTATGACCTTAACATCAGAACCGGGGTAAAGGACATC	2580
Qy	1441	GTTCCTCTGGAAGCTTTTGAAGACAAACAAAGTCTGTAGGACCCCTTTGCGAGGACGG	1500	Db	2521	GCGTATATGTTAAAGGCAATGTTATGACCTTAACATCAGAACCGGGGTAAAGGACATC	2580
Db	1441	GTTCCTCTGGAAGCTTTTGAAGACAAACAAAGTCTGTAGGACCCCTTTGCGAGGACGG	1500	Qy	2581	ACCAAGGGTGGCCCATCAGCTACTCCACTTATGGCAAGTTTCTTTCGCCGACGTTGCTG	2640
Qy	1501	AACCCCTCTGGAAGCTTTTGAAGACAAACAAAGTCTGTAGGACCCCTTTGCGAGGACGG	1560	Db	2581	ACCAAGGGTGGCCCATCAGCTACTCCACTTATGGCAAGTTTCTTTCGCCGACGTTGCTG	2640
Db	1501	AACCCCTCTGGAAGCTTTTGAAGACAAACAAAGTCTGTAGGACCCCTTTGCGAGGACGG	1560	Qy	2641	TCCTGGGGCGCCTATGACATCATATATGATGAGTGCCACTCAACTGATCGACCACT	2700
Qy	1561	GCAAGGCGGCAACACCCAGTCCACAGTGTGAGTTGATGTTGTAAGTGTGGAAGAGTCAA	1620	Db	2641	TCCTGGGGCGCCTATGACATCATATATGATGAGTGCCACTCAACTGATCGACCACT	2700
Db	1561	GCAAGGCGGCAACACCCAGTCCACAGTGTGAGTTGATGTTGTAAGTGTGGAAGAGTCAA	1620	Qy	2701	ATCTGGGCGATCGGACAGCTCTGGAACCAAGCGGACCGCTGGAGCGGACTCGTCTG	2760
Qy	1621	TGGCTCTCTCAAGCGTATTCAACAAAGGGCTGAAAGTGCAGGATGCCAGAGGTACCCATTGT	1680	Db	2701	ATCTGGGCGATCGGACAGCTCTGGAACCAAGCGGACCGCTGGAGCGGACTCGTCTG	2760
Db	1621	TGGCTCTCTCAAGCGTATTCAACAAAGGGCTGAAAGTGCAGGATGCCAGAGGTACCCATTGT	1680	Qy	2761	CTCGCACCGCTACGCTCGGATCGGTACCGTCCACATCCATCAACATCGAGGAGGTG	2820
Qy	1681	ATGGATCTGATCTGGGCGCTCGGTGACATGCTTTACATGTTTGTAGTTCGAGTTAAAA	1740	Db	2761	CTCGCACCGCTACGCTCGGATCGGTACCGTCCACATCCATCAACATCGAGGAGGTG	2820
Db	1681	ATGGATCTGATCTGGGCGCTCGGTGACATGCTTTACATGTTTGTAGTTCGAGTTAAAA	1740	Qy	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880
Qy	1741	AAGCTCTAGCCGCCCGAACACAGGAGCGTGTGTTTCTTTGAAAAACACGATATACC	1800	Db	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880
Db	1741	AAGCTCTAGCCGCCCGAACACAGGAGCGTGTGTTTCTTTGAAAAACACGATATACC	1800	Qy	2881	AAGGGGGAGGACCTCTATTTCTGCGCATTCAGAGAGAAATCTGATGAGCTCGCGCG	2940
Qy	1801	ATGGCGCTATTAGCGCTACTCCCAACAGCGAGCGCTTACCTTGGCTGATCATCACT	1860	Db	2881	AAGGGGGAGGACCTCTATTTCTGCGCATTCAGAGAGAAATCTGATGAGCTCGCGCG	2940
Db	1801	ATGGCGCTATTAGCGCTACTCCCAACAGCGAGCGCTTACCTTGGCTGATCATCACT	1860	Qy	2941	AAGTGTCTCGGCTCGGACTCAATGCTGTAGCATATTTACCGGGGCTTGTATGCTGTC	3000
Qy	1861	AGCTTCAGCGCGGACAGAAACAGGTCCAGGGGAGGTCCAAAGTGTCTCCACGCA	1920	Db	2941	AAGTGTCTCGGCTCGGACTCAATGCTGTAGCATATTTACCGGGGCTTGTATGCTGTC	3000
Db	1861	AGCTTCAGCGCGGACAGAAACAGGTCCAGGGGAGGTCCAAAGTGTCTCCACGCA	1920	Qy	3001	ATACCAACTAGCGGAGAGCTCATTTGCTGTAGCAACGAGCGCTTAATGACGGGCTTACC	3060
Qy	1921	ACAAATCTTCTGCGGACCTGCTCAATGGCGGTGTGTTGGACTGTCTATCATGGTGC	1980	Db	3001	ATACCAACTAGCGGAGAGCTCATTTGCTGTAGCAACGAGCGCTTAATGACGGGCTTACC	3060
Db	1921	ACAAATCTTCTGCGGACCTGCTCAATGGCGGTGTGTTGGACTGTCTATCATGGTGC	1980	Qy	3061	GGCGATTTGAGTCACTGAGTGCATTCATGCTGTCACCCAGACAGTCACTTCACTG	3120
Qy	1981	GGCTCAAGAGCCCTTGGCGGCGCAAGGGCCCAATCACCCAAATGTACACCAATGTGAC	2040	Db	3061	GGCGATTTGAGTCACTGAGTGCATTCATGCTGTCACCCAGACAGTCACTTCACTG	3120
Db	1981	GGCTCAAGAGCCCTTGGCGGCGCAAGGGCCCAATCACCCAAATGTACACCAATGTGAC	2040	Qy	3121	CTGACCCGACTTCACTTACCATTTGACGACGACCGTCCACAGAGCGGTGTCACTG	3180
Qy	2041	CAGGACCTCTGCTGGCTGACAGCGCCCTCGGCGGCTTCTTGGACACCATGACCTGCG	2100	Db	3121	CTGACCCGACTTCACTTACCATTTGACGACGACCGTCCACAGAGCGGTGTCACTG	3180
Db	2041	CAGGACCTCTGCTGGCTGACAGCGCCCTCGGCGGCTTCTTGGACACCATGACCTGCG	2100	Qy	3181	CAGGGGCGAGGACGACTGTTAGGGGAGGATGAGGCAATTTACAGTTTGTGACTCCAGGA	3240
Qy	2101	GGCAGCTCGGACCTTTACTTGTGTCAGAGGACATGCCGATGTCATTCGGGTGCGCGGG	2160	Db	3181	CAGGGGCGAGGACGACTGTTAGGGGAGGATGAGGCAATTTACAGTTTGTGACTCCAGGA	3240
Db	2101	GGCAGCTCGGACCTTTACTTGTGTCAGAGGACATGCCGATGTCATTCGGGTGCGCGGG	2160				

QY 3241 GAACGGCCCTCGGCGATGTTTCGATCTCTGCTGAGTGTATGACGGGGCTGT 3300  
Db 3241 GAACGGCCCTCGGCGATGTTTCGATCTCTGCTGAGTGTATGACGGGGCTGT 3300  
QY 3301 GCTTGCTAGAGCTACGCGCCGCGAGACCTCAGTTAGTTGGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGCTAGAGCTACGCGCCGCGAGACCTCAGTTAGTTGGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGGTTCGCCGTCTGCCAGGACCACTTGGAGTCTTGGGAGAGCTCTTTTACAGCCCTC 3420  
Db 3361 CCAGGGTTCGCCGTCTGCCAGGACCACTTGGAGTCTTGGGAGAGCTCTTTTACAGCCCTC 3420  
QY 3421 ACCACATAGACGCCCATTTCTTGTCCGAGACTAAGCAGGAGGACAACTTCCCTTAC 3480  
Db 3421 ACCACATAGACGCCCATTTCTTGTCCGAGACTAAGCAGGAGGACAACTTCCCTTAC 3480  
QY 3481 CTGGTAGCATACGAGCTACGGTGTGCGCCAGGGCTCAGGCTCCACTCGTGGGAC 3540  
Db 3481 CTGGTAGCATACGAGCTACGGTGTGCGCCAGGGCTCAGGCTCCACTCGTGGGAC 3540  
QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAGGGCCCAACGCCCTGCTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAGGGCCCAACGCCCTGCTG 3600  
QY 3601 TATAGGCTGGGAGCGGTTCAAACGAGGTTACTACCACACACCCCATACCAATACATC 3660  
Db 3601 TATAGGCTGGGAGCGGTTCAAACGAGGTTACTACCACACACCCCATACCAATACATC 3660  
QY 3661 ATGGCATCATGTCGGCTGACCTGAGGTGCTCAGAGCACTTGGGTCTGGTAGGCGGA 3720  
Db 3661 ATGGCATCATGTCGGCTGACCTGAGGTGCTCAGAGCACTTGGGTCTGGTAGGCGGA 3720  
QY 3721 GTCCTAGCAGCTCTGGCCGCGTATTGCTGCTCAACAGCAGCGGTGCTATTTGGGCGAGG 3780  
Db 3721 GTCCTAGCAGCTCTGGCCGCGTATTGCTGCTCAACAGCAGCGGTGCTATTTGGGCGAGG 3780  
QY 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTTCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTTCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGTGGCTCTACACTCTCCTTTACATGCAACAGGGAATGACCTCGCC 3900  
Db 3841 GATGAGATGGAAGTGGCTCTACACTCTCCTTTACATGCAACAGGGAATGACCTCGCC 3900  
QY 3901 GAAACAATTCAAACAGAGGCAATCGGGTGTGCTGCAACAGCCACCAAGCAGCGAGGCT 3960  
Db 3901 GAAACAATTCAAACAGAGGCAATCGGGTGTGCTGCAACAGCCACCAAGCAGCGAGGCT 3960  
QY 3961 GCTGCTCCCGTGTGGAAATCCAGTGGCGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCCGTGTGGAAATCCAGTGGCGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
QY 4021 TGGAAATTCATCAGCGGGATACAAATATTAGAGGCTTTGCTGCTGCTGCAACCCCC 4080  
Db 4021 TGGAAATTCATCAGCGGGATACAAATATTAGAGGCTTTGCTGCTGCTGCAACCCCC 4080  
QY 4081 GCGATAGCATCACTGATGGCAATTCACAGCTCTATCACCAGCGGCTCACCACCCACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCAATTCACAGCTCTATCACCAGCGGCTCACCACCCACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTGCTTCTGAGCGCCCGCATCGCTGAGCGGCTTGGGAGCATAGGCTTTGGG 4260  
Db 4201 GCTTCTGCTTCTGAGCGCCCGCATCGCTGAGCGGCTTGGGAGCATAGGCTTTGGG 4260  
QY 4261 AAGGTGCTTGTGATATTTTGGCAGGTATGAGCAGGGGTGGCAGGCGCGCTGCTGGCC 4320  
Db 4261 AAGGTGCTTGTGATATTTTGGCAGGTATGAGCAGGGGTGGCAGGCGCGCTGCTGGCC 4320  
QY 4321 TTTAAGGTCAAGCGCGGAGATGCCCTCCACCGAGGACCTGTTTAACTCTCCTGCT 4380

Db 4321 TTTAAGGTCAAGCGCGGAGATGCCCTCCACGAGGACCTGTTTAACTCTCCTGCT 4380  
QY 4381 ATCTCTCTCCCTGGCGCCCTAGTCTGTCGGGTCTGTCGGCAGCATCTGCTCGGCAC 4440  
Db 4381 ATCTCTCTCCCTGGCGCCCTAGTCTGTCGGGTCTGTCGGCAGCATCTGCTCGGCAC 4440  
QY 4441 GTGGGCCCAAGGGAGGGGCTGTGTCAGTGGATGAACCGGCTGATAGCTTTCGTTTCGGG 4500  
Db 4441 GTGGGCCCAAGGGAGGGGCTGTGTCAGTGGATGAACCGGCTGATAGCTTTCGTTTCGGG 4500  
QY 4501 GGTAAACCAAGTCTCCCCACGCACTATGTCCTGAGAGCGCTGCACGCTGTCAC 4560  
Db 4501 GGTAAACCAAGTCTCCCCACGCACTATGTCCTGAGAGCGCTGCACGCTGTCAC 4560  
QY 4561 CAGATCTCTCTGATGCTTACCATCACTCAGCTGCTGAAGAGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCTCTCTGATGCTTACCATCACTCAGCTGCTGAAGAGCTTCAACAGTGGATCAAC 4620  
QY 4621 GAGGACTCTCTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
Db 4621 GAGGACTCTCTCCACGCCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
QY 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTCGCGGATTTGCGGGA 4740  
Db 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTCGCGGATTTGCGGGA 4740  
QY 4741 GTCCCTCTTCTCATGTCACGCTGAGGCTTCAAGGAGTCTGGCGGGGACGGCATCATG 4800  
Db 4741 GTCCCTCTTCTCATGTCACGCTGAGGCTTCAAGGAGTCTGGCGGGGACGGCATCATG 4800  
QY 4801 CAAACCACTCTCCATGTCGAGCAGATCAACCGGACATGTAAGAAACGGTTCATGAGG 4860  
Db 4801 CAAACCACTCTCCATGTCGAGCAGATCAACCGGACATGTAAGAAACGGTTCATGAGG 4860  
QY 4861 ATCTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCATTTAACCGGTAC 4920  
Db 4861 ATCTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTTCCCATTTAACCGGTAC 4920  
QY 4921 ACCACGGCCCTTCACGCCCTCCCCGGCGCAATTTATTTAGGGCGCTGTGGCGGGT 4980  
Db 4921 ACCACGGCCCTTCACGCCCTCCCCGGCGCAATTTATTTAGGGCGCTGTGGCGGGT 4980  
QY 4981 GCTGCTAGGAGTACGTCGAGGTTTACCGGCTGGGGATTTCCACATACGTCAGCGGATG 5040  
Db 4981 GCTGCTAGGAGTACGTCGAGGTTTACCGGCTGGGGATTTCCACATACGTCAGCGGATG 5040  
QY 5041 ACCACTGAACAGTAAAGTCCCGTGTAGGTTCCGGCCCCCGAAATTTCTTCAAGAAAGT 5100  
Db 5041 ACCACTGAACAGTAAAGTCCCGTGTAGGTTCCGGCCCCCGAAATTTCTTCAAGAAAGT 5100  
QY 5101 GATGGGTCGGTTCACAGGTACGTCAGCGTCCAGCGTGAACCCCTCTTACGGGAGGTC 5160  
Db 5101 GATGGGTCGGTTCACAGGTACGTCAGCGTCCAGCGTGAACCCCTCTTACGGGAGGTC 5160  
QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGTCCCATCGGAGGTC 5220  
Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGTCCCATCGGAGGTC 5220  
QY 5221 CCGGAGTACGAGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACGGCGGAGACG 5280  
Db 5221 CCGGAGTACGAGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACGGCGGAGACG 5280  
QY 5281 GCTAAGCGTACGCTCCAGGGGATCTCCCGCTCTTGGCCAGCTCATCAGTACGCCAG 5340  
Db 5281 GCTAAGCGTACGCTCCAGGGGATCTCCCGCTCTTGGCCAGCTCATCAGTACGCCAG 5340  
QY 5341 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCCGAGCGTGCAC 5400  
Db 5341 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCCGAGCGTGCAC 5400  
QY 5401 CTCATCGAGGCAACCTCTCTGCGGCGAGGATGGCGGGAACATCACCGCGCTGGAG 5460



QY 7621 TTATCAGCTGGTTCCTTCTGCTGTTACAGCGGGGAGACATATATACACAGCCTGTCTCT 7680  
Db |||||  
QY 7621 TTATCAGCTGGTTCCTTCTGCTGTTACAGCGGGGAGACATATATACAGCCTGTCTCT 7680  
Db |||||  
QY 7681 GCGGACCCCGCTGGTTCATGTGGTCCCTACTCTTCTGTAGGGTAGGCATCTAT 7740  
Db |||||  
QY 7741 CTACTCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTT 7800  
Db |||||  
QY 7741 CTACTCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTT 7800  
Db |||||  
QY 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7861 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7921 TAGCTGTGAAGGTCCTGAGCGCTTGACTGCAGAGAGTGTGATCTAGCCTCTCTGC 7980  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||  
RESULT 11  
AAL47281  
ID AAL47281 standard; DNA; 7992 BP.  
XX  
AC AAL47281;  
XX  
DT 30-AUG-2002 (first entry)  
XX  
DE Hepatitis C virus sub-genomic replicon recombinant clone HCVR24.  
XX  
KW Hepatitis; HCV; core-neo; NS3 proteinase/helicase; vaccine; diagnosis;  
KW virucide; hepatotropic; gene therapy; anti-viral; gene; ds.  
XX  
OS Hepatitis C virus.  
XX  
FN WO200238793-A2.  
XX  
PD 16-MAY-2002.  
XX  
PF 02-NOV-2001; 2001WO-US046350.  
XX  
PR 07-NOV-2000; 2000US-0245866P.  
XX  
PA (ANAD-) ANADYS PHARM INC.  
XX  
PI Bichko V;  
XX  
XX WPI; 2002-490082/52.  
XX  
PT Novel nucleic acid encoding replication competent recombinant hepatitis C  
PT virus genome useful for screening anti-hepatitis C virus therapeutics and  
PT for vaccine development.  
XX  
PS Claim 11; Page 70-75; 85pp; English.  
XX  
CC The present invention provides protein and coding sequences from  
CC Hepatitis C virus (HCV), comprising all or part of the HCV genome and  
CC able to replicate efficiently when transfected into a susceptible cell  
CC line without reducing the growth rate of the cell line by more than 10  
CC fold. The sequences are useful for screening for anti-HCV therapeutics,  
CC for detecting antibodies to HCV in a biological sample such as blood,  
CC serum, plasma, blood cells, lymphocytes, or liver cells from a subject,  
CC for deriving authentic HCV components such as replication-complement non-  
CC infectious, replication-defective infection-component, and replication-

CC defective non-infectious HCV, in gene therapy or gene vaccination  
CC targeted to hepatic tissue for treating an animal infected or susceptible  
CC to HCV infection and for studying HCV infection and propagation. The  
CC present sequence is a clone of a fragment of the HCV genome designated  
CC HCVR24  
XX  
SQ Sequence 7992 BP; 1648 A; 2369 C; 2242 G; 1733 T; 0 U; 0 Other;  
Query Match 100.0%; Score 7985.8; DB 6; Length 7992;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 7987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 GCCAGCCCCGATTGGGGCGACACTCCACCATAGATCATCCCTCTGTGAGAACTACTG 60  
Db |||||  
QY 1 GCCAGCCCCGATTGGGGCGACACTCCACCATAGATCATCCCTCTGTGAGAACTACTG 60  
Db |||||  
QY 61 TCTTCACGACGAAACGCTTAGCCATGGCGTTAGTATGAGTGTCTGAGCCTCAGGAC 120  
Db |||||  
QY 61 TCTTCACGACGAAACGCTTAGCCATGGCGTTAGTATGAGTGTCTGAGCCTCAGGAC 120  
Db |||||  
QY 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGGGAAACCGGTGAGTACACCGGAATTGCCAG 180  
Db |||||  
QY 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGGGAAACCGGTGAGTACACCGGAATTGCCAG 180  
Db |||||  
QY 181 GACGACCGGTCCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGGCGTCCCCCC 240  
Db |||||  
QY 181 GACGACCGGTCCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGGCGTCCCCCC 240  
Db |||||  
QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGTCCGAAAGACCTTGGTACTGCTGATAGG 300  
Db |||||  
QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGTCCGAAAGACCTTGGTACTGCTGATAGG 300  
Db |||||  
QY 301 GTGCTTCGAGTGCCTCCGGAGGTCTCGTAGACCGTCACCATGAGCAGATCTTAAC 360  
Db |||||  
QY 301 GTGCTTCGAGTGCCTCCGGAGGTCTCGTAGACCGTCACCATGAGCAGATCTTAAC 360  
Db |||||  
QY 361 CTAAAGAAACCAAGGGCGGCCATGATTGAACAGATGGATTGCACGAGTCTC 420  
Db |||||  
QY 361 CTAAAGAAACCAAGGGCGGCCATGATTGAACAGATGGATTGCACGAGTCTC 420  
Db |||||  
QY 421 CGCCGCTTGGTGGAGAGCTATTGCGCTATGACTGGGCAACACAAATCGGCTGCT 480  
Db |||||  
QY 421 CGCCGCTTGGTGGAGAGCTATTGCGCTATGACTGGGCAACACAAATCGGCTGCT 480  
Db |||||  
QY 481 CTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGCCGGTCTTTTGTCAAGACCG 540  
Db |||||  
QY 481 CTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGCCGGTCTTTTGTCAAGACCG 540  
Db |||||  
QY 541 ACTGTCCGCTGCCCTGAATGAATGAACTGAGGACGAGCGCGCTATCGTGGTGGCCA 600  
Db |||||  
QY 541 ACTGTCCGCTGCCCTGAATGAATGAACTGAGGACGAGCGCGCTATCGTGGTGGCCA 600  
Db |||||  
QY 601 CGACGGCGCTTCTTGGCAGCTGTCTGAGCTTGTCTGAGCGGGAAGGACTGGC 660  
Db |||||  
QY 601 CGACGGCGCTTCTTGGCAGCTGTCTGAGCTTGTCTGAGCGGGAAGGACTGGC 660  
Db |||||  
QY 661 TGCTATTGGCGAAGTCCCGGGCAGGATCTCTGTCTATCTACCTTGTCTTCCCGCAGA 720  
Db |||||  
QY 661 TGCTATTGGCGAAGTCCCGGGCAGGATCTCTGTCTATCTACCTTGTCTTCCCGCAGA 720  
Db |||||  
QY 721 AAGTATCCATATGCTGATGCAATGGCGGCTGCATACGCTTGTATCCGGTACTCTGCC 780  
Db |||||  
QY 781 CATTTCGACCAACAGGAAACATCGATCGAGCAGCAGTACTCGGATGGAAGCGGTC 840  
Db |||||  
QY 781 CATTTCGACCAACAGGAAACATCGATCGAGCAGCAGTACTCGGATGGAAGCGGTC 840  
Db |||||  
QY 841 TTGTTCGATCAGATGATCTGGACGAGGAGCATCAGGGCTCGCGCAGCGGAACTGTCG 900  
Db |||||  
QY 841 TTGTTCGATCAGATGATCTGGACGAGGAGCATCAGGGCTCGCGCAGCGGAACTGTCG 900  
Db |||||  
QY 901 CCAGGCTCAAGGCGCGCATGTCGCCGAGGAGATCTCGTGTGACCATGCGATGCT 960  
Db |||||

Db	901		CCAGGCTCAAGGCGCGATGCGCGAGGATCTCGTGACCATGGCGATGCGCT	960	Db	1981	GGCTCAAGAGCCCTTGCGCGCCCAAGAGGCGCAATCACCACCAATGTACACCAATGTGGAC	2040
Qy	961		GCTTGCCGAATATCATGTGGAATAATGCGCGCTTTCTGAGATTCATCGACTGTGGCGGC	1020	Qy	2041	CAGGACCTCGTGCCTGGCAAGCGCCCCCGGGGGCGGTTCCTTTGACACCAATGACCTGCG	2100
Db	961			1020	Db	2041	CAGGACCTCGTGCCTGGCAAGCGCCCCCGGGGGCGGTTCCTTTGACACCAATGACCTGCG	2100
Qy	1021		TGGTGTGGCGGACCGCTATCAGGACATAGCGTTCGGTACCGGTGATATTCGTGAAGAGC	1080	Qy	2101	GGCAGCTCGGACCTTTACTTGGTCAAGAGGATGCGGATGTCAATTCGGTGGCGGGCGG	2160
Db	1021			1080	Db	2101	GGCAGCTCGGACCTTTACTTGGTCAAGAGGATGCGGATGTCAATTCGGTGGCGGGCGG	2160
Qy	1081		TTGGCGGGAATGGGCTGACCGCTTCCTGCTGCTTACGGTATCGCGCTCCCGATTCGC	1140	Qy	2161	GGGACACAGGCGGAGCTACTCTCCCGAGGCGCGTCTCTACTTGAAGGGCTCTTCG	2220
Db	1081			1140	Db	2161	GGGACACAGGCGGAGCTACTCTCCCGAGGCGCGTCTCTACTTGAAGGGCTCTTCG	2220
Qy	1141		AGCGATCGCCTTCATCGCCTTCCTGACGAGTTCCTCTGAGTTTAAACAGACCAACAG	1200	Qy	2221	GGCGTTCACATGCTCTGCGCCCTCGGGGACGCTGTGGGATCTTTGGGCTGCGGTGTC	2280
Db	1141			1200	Db	2221	GGCGTTCACATGCTCTGCGCCCTCGGGGACGCTGTGGGATCTTTGGGCTGCGGTGTC	2280
Qy	1201		GTTCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCTTAACGTACTGCG	1260	Qy	2281	ACCCGAGGGTTCGAAAGCGGTGAGCTTTGTACCGTCTGAGTCTATGGAAACCACTATG	2340
Db	1201			1260	Db	2281	ACCCGAGGGTTCGAAAGCGGTGAGCTTTGTACCGTCTGAGTCTATGGAAACCACTATG	2340
Qy	1261		CGAAGCGCTTGAATAAGCGCGGTGCGTTTGTCTATATGTTATTTTCCACCATATG	1320	Qy	2341	CGGTCCCGGTCTTCAAGGACAACTCGTCCCTCGCGCGTACCGGACACATTTCCAGGTG	2400
Db	1261			1320	Db	2341	CGGTCCCGGTCTTCAAGGACAACTCGTCCCTCGCGCGTACCGGACACATTTCCAGGTG	2400
Qy	1321		CGGTCTTTTGGAAATGAGGCGCGGAAACCTGGCCCTGTCTTTGACGAGCATTCCT	1380	Qy	2401	GCCCATCTACACGCCCTACTCTGAGCGGCAAGAGACATTAAGTGGCGGCTGGGTATGCA	2460
Db	1321			1380	Db	2401	GCCCATCTACACGCCCTACTCTGAGCGGCAAGAGACATTAAGTGGCGGCTGGGTATGCA	2460
Qy	1381		AGGGGTCTTTCCTCTGCGCAAGGAATGCAAGGTCTGTTGAATCTGCTGAAGGAGCA	1440	Qy	2461	GCCAAAGGATTAAGGTCTTGTCTGAACCCGTCTGCGCGGCGGCGGCTGGGTATGCA	2520
Db	1381			1440	Db	2461	GCCAAAGGATTAAGGTCTTGTCTGAACCCGTCTGCGCGGCGGCGGCTGGGTATGCA	2520
Qy	1441		GTTCCTCTGGAAGCTTTGAGACAAACAACTCTGTAGCGACCTTTTCAGCGAGCGG	1500	Qy	2521	GGGTATCTGTAAGGACATGCTATCGACCTTAACATCAGAACCGGGGTAAAGACCATC	2580
Db	1441			1500	Db	2521	GGGTATCTGTAAGGACATGCTATCGACCTTAACATCAGAACCGGGGTAAAGACCATC	2580
Qy	1501		AACCCCGCTCTGCGACAGGTGCTCTGCGGCGCAAGAGCGAGTGTATAGATACACCT	1560	Qy	2581	ACCAAGGCTGCGCCCATCAGTACTCCACTTGGCAAGTTCCTTGGCGAGCGTGGTTCG	2640
Db	1501			1560	Db	2581	ACCAAGGCTGCGCCCATCAGTACTCCACTTGGCAAGTTCCTTGGCGAGCGTGGTTCG	2640
Qy	1561		GCAAGGCGGACAAACCCAGTGCACGTTGTGAGTTGATGATGTTGGAAAGAGTCAAA	1620	Qy	2641	TCCTGGGCGGCTTATGACATCATATATGATGAGTGCCACTCAACTGACTCGACCAT	2700
Db	1561			1620	Db	2641	TCCTGGGCGGCTTATGACATCATATATGATGAGTGCCACTCAACTGACTCGACCAT	2700
Qy	1621		TGGCTCTCTCAAGCGTATTTCAACAGGGGCTGAAGGATGCCAGAGGTACCCCATTTG	1680	Qy	2701	ATCCTGGGCTATCGGACAGTCTTGGACCAAGCGGAGAGGCTGGAGCGGACTCGTGTG	2760
Db	1621			1680	Db	2701	ATCCTGGGCTATCGGACAGTCTTGGACCAAGCGGAGAGGCTGGAGCGGACTCGTGTG	2760
Qy	1681		ATGGATCTCATCTGGGCGCTCGGTGCACATGCTTTACATGTTTATGTCGAGTTTAAA	1740	Qy	2761	CTGCGCACCGCTACGCTCGGGGATCGGTCCACCTGCGGCGGCGGCTGGAGAGGTG	2820
Db	1681			1740	Db	2761	CTGCGCACCGCTACGCTCGGGGATCGGTCCACCTGCGGCGGCGGCTGGAGAGGTG	2820
Qy	1741		AACGTCTAGGCGGCGGACCGTGGTTCCTTTGAAACACGATAATACC	1800	Qy	2821	GCTCTGTCAGCACTGGAGAAATCCCTTTATGGCAAGAGCAATCCCATCGAGACCATC	2880
Db	1741			1800	Db	2821	GCTCTGTCAGCACTGGAGAAATCCCTTTATGGCAAGAGCAATCCCATCGAGACCATC	2880
Qy	1801		ATGGCGCTATTAGGCGCTACTCCCAACAGACGAGGCGCTTGGTGTGATCATCACT	1860	Qy	2881	AAGGGGGGAGGACCTCATTTCTGCGATTTCCAAAGAGAAATGTGATGAGCTCGCGCG	2940
Db	1801			1860	Db	2881	AAGGGGGGAGGACCTCATTTCTGCGATTTCCAAAGAGAAATGTGATGAGCTCGCGCG	2940
Qy	1861		AGCTTCACAGGCGGACAGGAAACAGGTTCGAGGGGAGGTCCAAAGTGGTTCACCGCA	1920	Qy	2941	AAGCTGTCCGCGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTGTATGTCCTC	3000
Db	1861			1920	Db	2941	AAGCTGTCCGCGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTGTATGTCCTC	3000
Qy	1921		ACACAATCTTCTGGCGACTGGTGTGATGAGTGTCTATCATGCTGTC	1980	Qy	3001	ATACCAACTAGCGGAGAGCTCATTTGTCTGAGCAAGGAGCTTAAATGAGGGCTTTACC	3060
Db	1921			1980	Db	3001	ATACCAACTAGCGGAGAGCTCATTTGTCTGAGCAAGGAGCTTAAATGAGGGCTTTACC	3060
Qy	1981		GGCTCAAGAGCCCTTGCGCGCCCAAGGCGCGCAATCACCACCAATGTACACCAATGTGGAC	2040	Qy	3061	GGCGATTTGCACTGAGTCAATGCTGCAATATGTCACCCAGACAGTGCAGTTCAGC	3120
Db	1981			2040	Db	3061	GGCGATTTGCACTGAGTCAATGCTGCAATATGTCACCCAGACAGTGCAGTTCAGC	3120



QY	3121	CTGGACCCGACCTTACCATTTAGACGACGACCGTGCCTCCAAAGACGGGTGTACGCTCG	3180
Db	3121	CTGGACCCGACCTTACCATTTAGACGACGACCGTGCCTCCAAAGACGGGTGTACGCTCG	3180
QY	3181	CAGCGGCAGGACGAGACTGTGTAGGGGACGATGGGCATTTACAGGTTTGTGACTCCAGGA	3240
Db	3181	CAGCGGCAGGACGAGACTGTGTAGGGGACGATGGGCATTTACAGGTTTGTGACTCCAGGA	3240
QY	3241	GAA CGGCGCTCGGGCATGTTCATTCCTCGGTTCTGTGCGAGTGTATGACGGGCTGT	3300
Db	3241	GAA CGGCGCTCGGGCATGTTCATTCCTCGGTTCTGTGCGAGTGTATGACGGGCTGT	3300
QY	3301	GCTTGGTACGAGCTCACGCCGCCGAGACCTCAGTTAGTTTCGGGAGGCTCTTTACAGGCTTC	3360
Db	3301	GCTTGGTACGAGCTCACGCCGCCGAGACCTCAGTTAGTTTCGGGAGGCTCTTTACAGGCTTC	3360
QY	3361	CCAGGGTTGCCGCTCTGCCAGGACCAATCTGGAGTTCTGGGAGGCTCTTTACAGGCTTC	3420
Db	3361	CCAGGGTTGCCGCTCTGCCAGGACCAATCTGGAGTTCTGGGAGGCTCTTTACAGGCTTC	3420
QY	3421	ACCCACATAGACGCCCATTTCTGTGCCAGACTAAGCAGGACGAGACAACCTTCCCTTAC	3480
Db	3421	ACCCACATAGACGCCCATTTCTGTGCCAGACTAAGCAGGACGAGACAACCTTCCCTTAC	3480
QY	3481	CTGGTAGCATACGAGCTACGGTGTGGCCAGGGCTCAGGCTCCACCTTCCATCTGTTGGAC	3540
Db	3481	CTGGTAGCATACGAGCTACGGTGTGGCCAGGGCTCAGGCTCCACCTTCCATCTGTTGGAC	3540
QY	3541	CAAATGTGGAAGTGTCTCATACGGCTAAGCCTACGTGTGCA CGGCGCAACGCCCTGCTG	3600
Db	3541	CAAATGTGGAAGTGTCTCATACGGCTAAGCCTACGTGTGCA CGGCGCAACGCCCTGCTG	3600
QY	3601	TATAGCTGGAGCCGCTTCAAAACGAGGTTACTACACACACCCCAATACCAATACATC	3660
Db	3601	TATAGCTGGAGCCGCTTCAAAACGAGGTTACTACACACACCCCAATACCAATACATC	3660
QY	3661	ATGGCATGATGTCCGCTACCTTGGAGTGTGTGCA CGAGCACCTGGGTGCTGTAGGCGGA	3720
Db	3661	ATGGCATGATGTCCGCTACCTTGGAGTGTGTGCA CGAGCACCTGGGTGCTGTAGGCGGA	3720
QY	3721	GTCTTAGCAGCTCTGGCGCGGTTATGGCTGACAAACGAGCGGTGCTTGTGGGCGAGG	3780
Db	3721	GTCTTAGCAGCTCTGGCGCGGTTATGGCTGACAAACGAGCGGTGCTTGTGGGCGAGG	3780
QY	3781	ATCATCTTGTCCGAAAGCGGCCATCATTTCCGACAGGAAATGCTTACCGGAGTTC	3840
Db	3781	ATCATCTTGTCCGAAAGCGGCCATCATTTCCGACAGGAAATGCTTACCGGAGTTC	3840
QY	3841	GATGAGATGGAGATGCGCCTCACACCTTCCCTTACATCGAAACAGGGAATGCACTCGCC	3900
Db	3841	GATGAGATGGAGATGCGCCTCACACCTTCCCTTACATCGAAACAGGGAATGCACTCGCC	3900
QY	3901	CAACAATTCAAACAGAGCAATCGGTTGCTGCAAAACGAGCAACGAGCGGAGGCT	3960
Db	3901	GAA CAATTCAAACAGAGCAATCGGTTGCTGCAAAACGAGCAACGAGCGGAGGCT	3960
QY	3961	GCTGCTCCGTTGGTGAATCCAAGTGGCGGACCTCGAAGACCTTCTGGGCGAGCATATG	4020
Db	3961	GCTGCTCCGTTGGTGAATCCAAGTGGCGGACCTCGAAGACCTTCTGGGCGAGCATATG	4020
QY	4021	TGGAAATTCATCAGCGGATACAATATTAGCAGGCTTGTCCACTCTGCTGGCAACCC	4080
Db	4021	TGGAAATTCATCAGCGGATACAATATTAGCAGGCTTGTCCACTCTGCTGGCAACCC	4080
QY	4081	GCATAGCATCACTGATGGCATTTACAGGCTTATCACAGCGGCTCACACCCCAACAT	4140
Db	4081	GCATAGCATCACTGATGGCATTTACAGGCTTATCACAGCGGCTCACACCCCAACAT	4140
QY	4141	ACCTCTCTGTTTAAACATCTCTGGGGGATGGGTGGCGGCCCACTTGTCTCTCCAGCGCT	4200
Db	4141	ACCTCTCTGTTTAAACATCTCTGGGGGATGGGTGGCGGCCCACTTGTCTCTCCAGCGCT	4200

QY	4201	GCTTCTGCTTTT	CGTAGGGCCCGC	CAATCGCTGAGCGGCTGTTTGGCAGCATAGACCTTGG	4260
Db	4201	GCTTCTGCTTTT	CGTAGGGCCCGC	CAATCGCTGAGCGGCTGTTTGGCAGCATAGACCTTGG	4260
QY	4261	AAGTGTCTTGGAT	TATTTTGGCAGTTAT	TGGAGCAGGGGTGGCAGCGCGCTCGTGGCC	4320
Db	4261	AAGTGTCTTGGAT	TATTTTGGCAGTTAT	TGGAGCAGGGGTGGCAGCGCGCTCGTGGCC	4320
QY	4321	TTTAAGGTCAATG	AGCGGAGATGCCCTCC	ACCGAGACCTGGTTAACTACTCCCTGCT	4380
Db	4321	TTTAAGGTCAATG	AGCGGAGATGCCCTCC	ACCGAGACCTGGTTAACTACTCCCTGCT	4380
QY	4381	ATCTCTCCCTGCG	CCCTAGTCTCGTGGGTCTGT	CGCAGCGATACTGCGTCGGCAC	4440
Db	4381	ATCTCTCCCTGCG	CCCTAGTCTCGTGGGTCTGT	CGCAGCGATACTGCGTCGGCAC	4440
QY	4441	GTGGGCCAGGGG	AGGGGCTGTGCTAGTGGATGA	ACCGGCTGTGCTTTCGGG	4500
Db	4441	GTGGGCCAGGGG	AGGGGCTGTGCTAGTGGATGA	ACCGGCTGTGCTTTCGGG	4500
QY	4501	GGTAACCACTCT	CCCCCAACGCACTAT	GTGCTCAGAGCGCTGACACGTGTCACT	4560
Db	4501	GGTAACCACTCT	CCCCCAACGCACTAT	GTGCTCAGAGCGCTGACACGTGTCACT	4560
QY	4561	CAGATCTCTTAG	CTTTAACAATCACTCAGCTGT	GAAAGGCTTCAACAGTGGATCAAC	4620
Db	4561	CAGATCTCTTAG	CTTTAACAATCACTCAGCTGT	GAAAGGCTTCAACAGTGGATCAAC	4620
QY	4621	GAGGACTGTCCA	CGCCATGCTCGGCTCTGCTA	GAGATGTTTGGATTTGGATATGC	4680
Db	4621	GAGGACTGTCCA	CGCCATGCTCGGCTCTGCTA	GAGATGTTTGGATTTGGATATGC	4680
QY	4681	ACGGTGTGA	CTGATTTCAAGACCTGGCTC	CAAGTCCAGTCCAGCGCATCATG	4740
Db	4681	ACGGTGTGA	CTGATTTCAAGACCTGGCTC	CAAGTCCAGTCCAGCGCATCATG	4740
QY	4741	GTCCCCCTTCT	TCATGTCAACGTGGGTACA	AGGAGTCTGGCGGGCGCAGCGCATCATG	4800
Db	4741	GTCCCCCTTCT	TCATGTCAACGTGGGTACA	AGGAGTCTGGCGGGCGCAGCGCATCATG	4800
QY	4801	CAAAACACCTGC	CACTGGAGCACAGATCA	CGGCATGTGAAACCGTTCATGAGG	4860
Db	4801	CAAAACACCTGC	CACTGGAGCACAGATCA	CGGCATGTGAAACCGTTCATGAGG	4860
QY	4861	ATCGTGGGCTT	AGGACCTGTAGTAAACGTGG	CATGGAACATTTCCCATTAACCGGTAC	4920
Db	4861	ATCGTGGGCTT	AGGACCTGTAGTAAACGTGG	CATGGAACATTTCCCATTAACCGGTAC	4920
QY	4921	ACCAAGGCGCTG	CAGCGCTCCCGGCGCAAA	TATTTAGGCGCTGTGGCGGGTG	4980
Db	4921	ACCAAGGCGCTG	CAGCGCTCCCGGCGCAAA	TATTTAGGCGCTGTGGCGGGTG	4980
QY	4981	GCTGTGAGGAGT	ACGTGGAGTTACGGGTGGGGG	ATTTCCACTAGTACGGGCATG	5040
Db	4981	GCTGTGAGGAGT	ACGTGGAGTTACGGGTGGGGG	ATTTCCACTAGTACGGGCATG	5040
QY	5041	ACCACTGACAA	CGTAAAGTGGCCGTGT	CAGGTTCCGGCCCCCGAATTTCTTACAGAAAGT	5100
Db	5041	ACCACTGACAA	CGTAAAGTGGCCGTGT	CAGGTTCCGGCCCCCGAATTTCTTACAGAAAGT	5100
QY	5101	GATGGGTGCGGT	TGCAAGGTACGCTC	ACGGTGCAAAACCCCTCTTACGGAGAGGTC	5160
Db	5101	GATGGGTGCGGT	TGCAAGGTACGCTC	ACGGTGCAAAACCCCTCTTACGGAGAGGTC	5160
QY	5161	ACATTCCTGTC	GGGCTCAATCAATCTCTG	TTGGGTACAGCTCCCATGTGAGCCCGAA	5220
Db	5161	ACATTCCTGTC	GGGCTCAATCAATCTCTG	TTGGGTACAGCTCCCATGTGAGCCCGAA	5220
QY	5221	CCGACGTAGAGT	GCTCACTTCCATGCT	CACCGACCCCTCCCATATTACGCGGAGACG	5280
Db	5221	CCGACGTAGAGT	GCTCACTTCCATGCT	CACCGACCCCTCCCATATTACGCGGAGACG	5280
QY	5281	GCTAAGGCTAGG	CTGGCCAGGGATCT	CCGCCCTCTTGGCCAGCTCATCAGCTAGCCAG	5340



```

QY 7501 GTCCGCGTAGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGCGCAAGTACCTCTTC 7560
Db 7501 GTCCGCGTAGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGCGCAAGTACCTCTTC 7560
QY 7561 AACTGGCAGTAAGACCAAGCTCAAATCTCAATCCCGCTCGCTCCAGTTGGAT 7620
Db 7561 AACTGGCAGTAAGACCAAGCTCAAATCTCAATCCCGCTCGCTCCAGTTGGAT 7620
QY 7621 TTATCCAGCTGCTGTTGCTGCTTACAGCGGGGAGACATATACAGCCCTGCTCTGT 7680
Db 7621 TTATCCAGCTGCTGTTGCTGCTTACAGCGGGGAGACATATACAGCCCTGCTCTGT 7680
QY 7681 GCCGACCCCGCTGTTGCTGCTTACAGCGGGGAGACATATACAGCCCTGCTCTGT 7740
Db 7681 GCCGACCCCGCTGTTGCTGCTTACAGCGGGGAGACATATACAGCCCTGCTCTGT 7740
QY 7741 CTACTCCCAACCGATGACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800
Db 7741 CTACTCCCAACCGATGACGGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800
QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
Db 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
QY 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
Db 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
QY 7921 TAGCTGTGAAGCTCCGTGAGCCGCTTGAAGTGCAGAGAGTCTGATCTGCGCTCTCTGC 7980
Db 7921 TAGCTGTGAAGCTCCGTGAGCCGCTTGAAGTGCAGAGAGTCTGATCTGCGCTCTCTGC 7980
QY 7981 AGATCAAGT 7989
Db 7981 AGATCAAGT 7989

RESULT 12
ID ABK91448 standard; DNA; 10690 BP.
AC ABK91448;
XX
XX
DT 15-NOV-2002 (first entry)
DE Hepatitis C virus vector construct pHCVNeo.17ml5.
KW HCV; ss; pHCVNeo.17m2; adaptive mutation; liver failure; cirrhosis;
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;
KW internal ribosome entry site; IRES; NS5A; NS5A; HCV replication; mutant.
XX
OS Hepatitis C virus.
OS Encephalomyocarditis virus.
OS Escherichia coli.
OS Enterobacteria phage T7.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT 5'UTR 1..341
FT CDS /*tag= a
FT 342..1181
FT /*tag= b
FT /product= "Core-neo fusion protein"
FT misc_signal 1190..1800
FT /*tag= c
FT 1801..1800
FT /label= IRES
FT /note= "Internal ribosome entry site from EMCV"
FT CDS 1801..7758
FT /*tag= d
FT /product= "Polyprotein"
FT /note= "Comprising NS3, NS4A, NS4B, NS5A and NS5B"
FT mutation replace(5320,G)

```

```

FT 3'UTR /*tag= e
FT 7759..7991 /*tag= f
FT misc_feature 7992..10690 /*tag= g
FT /*note= "Plasmid derived sequences"
XX
PN WO200259321-A2.
XX
PD 01-AUG-2002.
XX
PF 16-JAN-2002; 2002WO-EP000526.
XX
PR 23-JAN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX Claim 16; Page: 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX ; (3) a recombinant cell produced by introducing into a human hepatoma
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicon and HCV
XX replicon enhanced cells are useful in studying HCV replication and
XX expression, and HCV and host cell interactions, producing HCV RNA and
XX proteins, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is an HCV based vector
XX pHCVNeo.17 mutant of the invention. Note: The present sequence is not
XX shown in the specification but was created by the indexer using the HCV
XX vector sequence appearing as ABK91412 and the information in Claim 16
XX
XX Sequence 10690 BP; 2335 A; 3045 C; 2907 G; 2403 T; 0 U; 0 Other;
Query Match 100.0%; Score 7985.8; DB 6; Length 10690;
Best Local Similarity -100.0%; Pred. No. 0;
Matches 7987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTCTG 60
Db 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTCTG 60
QY 61 TCTTCAACGAGAAAGCGTCTAGCCATGGCTTGTAGTATGAGTGTCTGAGCTCCAGGAC 120
Db 61 TCTTCAACGAGAAAGCGTCTAGCCATGGCTTGTAGTATGAGTGTCTGAGCTCCAGGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATTCAG 180
Db 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATTCAG 180
QY 181 GACGACCCGGTCTCTTCTTGGATCAACCGGTCAATGCTGAGATTTGGGCGTCCCCC 240
Db 181 GACGACCCGGTCTCTTCTTGGATCAACCGGTCAATGCTGAGATTTGGGCGTCCCCC 240

```

Qy 241 GCGAGACTGTCAGCCGAGTAGTGTGGTCCGAAAGSCCTTGTGTAAGTCCCTGATAGG 300  
Db 241 GCGAGACTGTCAGCCGAGTAGTGTGGTCCGAAAGSCCTTGTGTAAGTCCCTGATAGG 300  
Qy 301 GTGCTTGGAGTGGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360  
Db 301 GTGCTTGGAGTGGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360  
Qy 361 CTCAAGAAAAACAAAGGGGCGCCATGATTAAGAAAGATGGATGCAAGCAGGTTCTC 420  
Db 361 CTCAAGAAAAACAAAGGGGCGCCATGATTAAGAAAGATGGATGCAAGCAGGTTCTC 420  
Qy 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAAGCAATCGGCTGCT 480  
Db 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAAGCAATCGGCTGCT 480  
Qy 481 CTGATGCGCGCTGTTCCGAGTGTGAGCGCAGGGGCGCCGGTCTTTTGTCAAGACCG 540  
Db 481 CTGATGCGCGCTGTTCCGAGTGTGAGCGCAGGGGCGCCGGTCTTTTGTCAAGACCG 540  
Qy 541 ACCTGTCGGGTGCCCTGAATGAATGCAAGGAGCAGGAGCGCGGCTATCGTGGCTGGCA 600  
Db 541 ACCTGTCGGGTGCCCTGAATGAATGCAAGGAGCAGGAGCGCGGCTATCGTGGCTGGCA 600  
Qy 601 CGACGGGCGTTCCTTGGCAGCTGTGCTCGAGTGTGCTCACTGAAGGGGAGGGAGTGGC 660  
Db 601 CGACGGGCGTTCCTTGGCAGCTGTGCTCGAGTGTGCTCACTGAAGGGGAGGGAGTGGC 660  
Qy 661 TGCATTGGGGGAGTGGCGGGGAGGATCTCTGTGCTCACTGCTTGTCTTCCGCGAGA 720  
Db 661 TGCATTGGGGGAGTGGCGGGGAGGATCTCTGTGCTCACTGCTTGTCTTCCGCGAGA 720  
Qy 721 AAGTATCCATCATGCTGATGCAATGCGGGGCTGCATACGCTGTGATCGGGTACCTGCC 780  
Db 721 AAGTATCCATCATGCTGATGCAATGCGGGGCTGCATACGCTGTGATCGGGTACCTGCC 780  
Qy 781 CATTCAGCACACNAGCGAACAATCGCATGAGCGGACGCTACTCCGATGGAAGCGGTG 840  
Db 781 CATTCAGCACACNAGCGAACAATCGCATGAGCGGACGCTACTCCGATGGAAGCGGTG 840  
Qy 841 TTGCTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGCAATGTTGC 900  
Db 841 TTGCTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCGCAATGTTGC 900  
Qy 901 CCAGGCTCAAGGCGCGCATGCCGAGGATCTCGTGTGACCCATGGCGATGSCCT 960  
Db 901 CCAGGCTCAAGGCGCGCATGCCGAGGATCTCGTGTGACCCATGGCGATGSCCT 960  
Qy 961 GCTTGGCGAATATCATGTTGGAATGGCGCTTTCTGGATTTCATCGACTGTGGCGGC 1020  
Db 961 GCTTGGCGAATATCATGTTGGAATGGCGCTTTCTGGATTTCATCGACTGTGGCGGC 1020  
Qy 1021 TGGGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCCGCTGATTTGCTGAAGAGC 1080  
Db 1021 TGGGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCCGCTGATTTGCTGAAGAGC 1080  
Qy 1081 TTGCGGCGAATGGGCTGACCGTTCCTCGTGTCTTACGCTATCGCGCTCCCGAATCGC 1140  
Db 1081 TTGCGGCGAATGGGCTGACCGTTCCTCGTGTCTTACGCTATCGCGCTCCCGAATCGC 1140  
Qy 1141 AGCGCATGCGCTTCTATCGCTTCTTCAGAGTTCCTTCGAGTTTAAACAGACCAACAG 1200  
Db 1141 AGCGCATGCGCTTCTATCGCTTCTTCAGAGTTCCTTCGAGTTTAAACAGACCAACAG 1200  
Qy 1201 GTTTCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCGCCCTTAAAGTTCGTCGGC 1260  
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCGCCCTTAAAGTTCGTCGGC 1260  
Qy 1261 CGAAGCGCTTGGAAATAGGCGGTGCTGTTGTTGTTATGTTTCCACCATATG 1320  
Db 1261 CGAAGCGCTTGGAAATAGGCGGTGCTGTTGTTGTTATGTTTCCACCATATG 1320  
Qy 1321 CGGTCTTTGGCAATGTGAGGGCCCGAAACCTGGCCCTGCTTCTTGACGAGCATTCCT 1380

Db 1321 CCCTCTTTTGGCAATGTGAGGGCCCGAAACCTGGCCCTGCTTCTTGACGAGCATTCCT 1380  
Qy 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAATCAAGGTCTGTTGAATGTCGTGAAGGAAGCA 1440  
Db 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAATCAAGGTCTGTTGAATGTCGTGAAGGAAGCA 1440  
Qy 1441 GTTCTCTGGAAGCTTCTTGAAGAACAACAAGCTGTGAGGACCCCTTTGAGGAGCGG 1500  
Db 1441 GTTCTCTGGAAGCTTCTTGAAGAACAACAAGCTGTGAGGACCCCTTTGAGGAGCGG 1500  
Qy 1501 AACCCCCACCTGGGACAGTGTCTCGCGCCAAAGACCAAGCTGTATAGATACACT 1560  
Db 1501 AACCCCCACCTGGGACAGTGTCTCGCGCCAAAGACCAAGCTGTATAGATACACT 1560  
Qy 1561 GCAAGGGGCGCACAAACCCAGTGCACCGTGTGAGTTGATAGTTGTGGAAGAGTCAAA 1620  
Db 1561 GCAAGGGGCGCACAAACCCAGTGCACCGTGTGAGTTGATAGTTGTGGAAGAGTCAAA 1620  
Qy 1621 TGGCTCTCTCAAGCGTATTCAACAAGGGGTGAAGGATGCCAGAGGTACCCATTGT 1680  
Db 1621 TGGCTCTCTCAAGCGTATTCAACAAGGGGTGAAGGATGCCAGAGGTACCCATTGT 1680  
Qy 1681 ATGGATCTGATCTGGGCTCGGTGCATGCTTTACATGTTTAGTCGAGGTAAAA 1740  
Db 1681 ATGGATCTGATCTGGGCTCGGTGCATGCTTTACATGTTTAGTCGAGGTAAAA 1740  
Qy 1681 ATGGATCTGATCTGGGCTCGGTGCATGCTTTACATGTTTAGTCGAGGTAAAA 1740  
Db 1681 ATGGATCTGATCTGGGCTCGGTGCATGCTTTACATGTTTAGTCGAGGTAAAA 1740  
Qy 1741 AACGCTTAGGCCCCCGAAACCAACCGGGAGTGTCTTTGAAAAACACGATAATAC 1800  
Db 1741 AACGCTTAGGCCCCCGAAACCAACCGGGAGTGTCTTTGAAAAACACGATAATAC 1800  
Qy 1801 ATGGGCTTATTACGGCTTATCCCAAGAGCGAGGCTACTTGGCTGCATCATCACT 1860  
Db 1801 ATGGGCTTATTACGGCTTATCCCAAGAGCGAGGCTACTTGGCTGCATCATCACT 1860  
Qy 1861 AGCTCAGAGCGGAGCAGGACAGGTGAGGGGAGGTCCAGTGTCTCCACCGCA 1920  
Db 1861 AGCTCAGAGCGGAGCAGGACAGGTGAGGGGAGGTCCAGTGTCTCCACCGCA 1920  
Qy 1921 ACACATCTTCTGCGGACCTGCGTCAATGSGGTGTGTTGAGTCTCTATCATGTGTC 1980  
Db 1921 ACACATCTTCTGCGGACCTGCGTCAATGSGGTGTGTTGAGTCTCTATCATGTGTC 1980  
Qy 1981 GGCTCAAGACCTTTCGCGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGAC 2040  
Db 1981 GGCTCAAGACCTTTCGCGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGAC 2040  
Qy 2041 CAGGACCTGTGCGTGGCAGCGCCCGGGGCGGTTCTTGACACCATGCACCTGC 2100  
Db 2041 CAGGACCTGTGCGTGGCAGCGCCCGGGGCGGTTCTTGACACCATGCACCTGC 2100  
Qy 2101 GGCAGCTCGGACCTTTTACTTGGTCAAGGATGCGGATGTCATTCGCGTGGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTTACTTGGTCAAGGATGCGGATGTCATTCGCGTGGCGCGG 2160  
Qy 2161 GGGCAGCAGGAGGAGCCTACTCTCCCGAGGCGCTCTCTACTTGAAGGGCTCTTCG 2220  
Db 2161 GGGCAGCAGGAGGAGCCTACTCTCCCGAGGCGCTCTCTACTTGAAGGGCTCTTCG 2220  
Qy 2221 GGGGCTTCACTGCTTCCCGCTGGGCGAGCTGTGGGCTCTTTCCGGGCTGCGGTGTC 2280  
Db 2221 GGGGCTTCACTGCTTCCCGCTGGGCGAGCTGTGGGCTCTTTCCGGGCTGCGGTGTC 2280  
Qy 2281 ACCCGAGGGTGGGAAGGGGTGGAATTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCCGAGGGTGGGAAGGGGTGGAATTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
Qy 2341 CGGTCCCGGTCTTCAAGCAAACTGTCCTCCCGGCGGTACCGCAGACATTCAGGTG 2400  
Db 2341 CGGTCCCGGTCTTCAAGCAAACTGTCCTCCCGGCGGTACCGCAGACATTCAGGTG 2400  
Qy 2401 GCGCATCTACAGCCCTTACTGTTAGGCGGAAGGACACTAAGGTGCGCGGTATGCA 2460







Db 6781 ACTAATCTTAAGGCGAGAACTCGCGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
QY 6841 ACCAGCTGGGTAAATACCTCTACATGTTACTTGAAGCGCGCTCGCGCTGTCGAGTGG 6900  
Db 6841 ACCAGCTGGGTAAATACCTCTACATGTTACTTGAAGCGCGCTCGCGCTGTCGAGTGG 6900  
QY 6901 AAGCTCCAGACTGACAGTGTCTGATGCGGAGACGACCTTGTCTGTTATCTGTGAAGC 6960  
Db 6901 AAGCTCCAGACTGACAGTGTCTGATGCGGAGACGACCTTGTCTGTTATCTGTGAAGC 6960  
QY 6961 GCGGGACCAAGAGGACGAGCGAGCTTACGGGCTTACGGGCTATGACTAGATAC 7020  
Db 6961 GCGGGACCAAGAGGACGAGCGAGCTTACGGGCTTACGGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTTGGGACCCGCCAAACAGAAATACGACTTGGAGTTGATACATCATG 7080  
Db 7021 TCTGCCCCCTTGGGACCCGCCAAACAGAAATACGACTTGGAGTTGATACATCATG 7080  
QY 7081 TCCTCCAATGTGTAGTGGCGGACGATGCTGCAAAAGGCTGACTTATCTCACCCGT 7140  
Db 7081 TCCTCCAATGTGTAGTGGCGGACGATGCTGCAAAAGGCTGACTTATCTCACCCGT 7140  
QY 7141 GACCCACACCCCTTGGCGGGCTGCTGGGAGACGCTAGACACACTCCAGTCAAT 7200  
Db 7141 GACCCACACCCCTTGGCGGGCTGCTGGGAGACGCTAGACACACTCCAGTCAAT 7200  
QY 7201 TCCTGGCTAGGCAACATCATGATGCGGCCACCTTGTGGCAAGGATGATCTGATG 7260  
Db 7201 TCCTGGCTAGGCAACATCATGATGCGGCCACCTTGTGGCAAGGATGATCTGATG 7260  
QY 7261 ACTCATTTCTCTCATTCCTCTAGCTCAGAAACAACTTGAAGCCCTAGATGTCAG 7320  
Db 7261 ACTCATTTCTCTCATTCCTCTAGCTCAGAAACAACTTGAAGCCCTAGATGTCAG 7320  
QY 7321 ATCTAGGGGCTGTTACTCCATTTAGGCACTTGACCTACCTCAGATCAATCAACGATC 7380  
Db 7321 ATCTAGGGGCTGTTACTCCATTTAGGCACTTGACCTACCTCAGATCAATCAACGATC 7380  
QY 7381 CATGGCCTTAGCGCATTTTCACTCCATGATTTACTCTCAGGTGAGATCAATAGGTTGCT 7440  
Db 7381 CATGGCCTTAGCGCATTTTCACTCCATGATTTACTCTCAGGTGAGATCAATAGGTTGCT 7440  
QY 7441 TCATGCTCAGAAACTTGGGTTACCGCCTTGGAGTCTGGAGACATCGGGCCAGAACT 7500  
Db 7441 TCATGCTCAGAAACTTGGGTTACCGCCTTGGAGTCTGGAGACATCGGGCCAGAACT 7500  
QY 7501 GTCGGGCTAGGCTAGTCTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTTTC 7560  
Db 7501 GTCGGGCTAGGCTAGTCTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTTTC 7560  
QY 7561 AACTGGGAGTAAGGACCAAGCTCAAACTCACTCCAACTCCGCTGCTGCCAGTTGGAT 7620  
Db 7561 AACTGGGAGTAAGGACCAAGCTCAAACTCACTCCAACTCCGCTGCTGCCAGTTGGAT 7620  
QY 7621 TTATCCAGCTGTTGTTGCTGTTTACGCGGGGAGACATATATACAGCCCTGTCTCGT 7680  
Db 7621 TTATCCAGCTGTTGTTGCTGTTTACGCGGGGAGACATATATACAGCCCTGTCTCGT 7680  
QY 7681 GCCCGACCCGCTGTTGTTGCTGTTTACCTACTTCTGAGGGTAGGCATCTAT 7740  
Db 7681 GCCCGACCCGCTGTTGTTGCTGTTTACCTACTTCTGAGGGTAGGCATCTAT 7740  
QY 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCAATAGGCCATCTCTTTT 7800  
Db 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCAATAGGCCATCTCTTTT 7800  
QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
QY 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920

QY 7921 TAGCTGTAAAGTCCGTGAGCGCTTGACTGCGAGAGTGTGATGACTGGCCTCTCTGC 7980  
Db 7921 TAGCTGTAAAGTCCGTGAGCGCTTGACTGCGAGAGTGTGATGACTGGCCTCTCTGC 7980  
QY 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989

RESULT 13  
ABK91435  
ID ABK91435 standard; DNA; 10690 BP.  
XX  
AC ABK91435;  
XX  
DT 15-NOV-2002 (first entry)  
XX  
DE Hepatitis C virus vector construct pHVNeo.17m2.  
XX  
KW HCV; ss; pHVNeo.17m2; adaptive mutation; liver failure; cirrhosis;  
KW hepatocellular carcinoma; NS3; NS5; encephalomyocarditis virus; EMCV;  
KW internal ribosome entry site; IRES; NS5A; HCV replication; mutant.  
XX  
OS Hepatitis C virus.  
OS Encephalomyocarditis virus.  
OS Escherichia coli.  
OS Enterobacteria phage T7.  
OS Synthetic.

Key Location/Qualifiers  
5'UTR 1..341  
CDS 342..1181  
FT /tag= a  
FT /tag= b  
FT /product= "Core-neo fusion protein"  
FT misc\_signal 1190..1800  
FT /tag= c  
FT /label= IRES  
FT /note= "Internal ribosome entry site from BCMV"  
FT CDS 1801..7758  
FT /tag= d  
FT /product= "Polyprotein"  
FT /note= "Comprising NS3, NS4A, NS4B, NS5A and NS5B"  
FT mutation replace(5243,C)  
FT 3'UTR 7759..7991  
FT /tag= e  
FT /tag= f  
FT misc\_feature 7992..10690  
FT /tag= g  
FT /note= "Plasmid derived sequences"

WO200259321-A2.  
PN  
XX  
PD 01-AUG-2002.  
XX  
PF 16-JAN-2002; 2002WO-EP000526.  
XX  
PR 23-JAN-2001; 2001US-0263479P.  
XX  
PA (RICE-) IST RICERHE BIOL MOLECOLARE ANGELETTI.  
XX  
PI De Francesco R, Migliaccio G, Paonessa G;  
XX WPI; 2002-599793/64.  
XX  
PT New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV  
PT NS5 encoding region, or encephalomyocarditis virus (EMCV) internal  
PT ribosome entry site (IRES) region, useful in studying HCV replication and  
PT expression.  
XX  
PS Claim 16; Page; 69pp; English.

The invention relates to nucleic acid molecules comprising altered HCV NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV) internal ribosome entry site (IRES) region coding for one or more NS3, NS5A, or EMCV IRES mutations, respectively. The location of the mutations are detailed in the specification. Also included are (1) an expression vector comprising a nucleotide sequence coding for the altered nucleic acids, which is transcriptionally coupled to an exogenous promoter; (2) a recombinant cell human hepatoma cell comprising the altered nucleic acids; (3) a recombinant cell produced by introducing into a human hepatoma cell the altered nucleic acids; (4) producing a functional HCV replicon; (5) replicon enhanced cell or which containing a functional HCV replicon; an HCV replicon enhanced cell made in the method; and (6) measuring the ability of a compound to affect HCV activity. The HCV replicon and HCV replicon enhanced cells are useful in studying HCV replication and expression, and HCV and host cell interactions, producing HCV RNA and proteins, and providing a system for measuring the ability of a compound to modulate one or more HCV activities e.g. to discover drugs which may treat HCV mediated diseases such as liver failure, cirrhosis and hepatocellular carcinoma. The present sequence is an HCV based vector pHCVNeo.17 mutant of the invention. Note: The present sequence is not shown in the specification but was created by the indexer using the HCV vector sequence appearing as ABK91412 and the information in Claim 16

XX Sequence 10690 BP; 2334 A; 3044 C; 2908 G; 2404 T; 0 U; 0 Other;

Query Match 100.0%; Score 7985.8; DE 6; Length 10690;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCACGCCCGGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 DB 1 GCACGCCCGGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60

QY 61 TCTTACGCGAGAAGCGTCTAGCATGGCTTGTAGTATGATGTCGTGAGCTCCAGGAC 120  
 DB 61 TCTTACGCGAGAAGCGTCTAGCATGGCTTGTAGTATGATGTCGTGAGCTCCAGGAC 120

QY 121 CCCCCCTCCCGGAGAGCCACTAGTGGTCTGCGAAACCGGTGATACACCGGAATTGCCAG 180  
 DB 121 CCCCCCTCCCGGAGAGCCACTAGTGGTCTGCGAAACCGGTGATACACCGGAATTGCCAG 180

QY 181 GACGACGGGTCTTTCTTGATCAACCCCTCAATGCCCTGGAGATTGGGGCGTGCCCC 240  
 DB 181 GACGACGGGTCTTTCTTGATCAACCCCTCAATGCCCTGGAGATTGGGGCGTGCCCC 240

QY 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGGTACTGCTGATAGG 300  
 DB 241 GCGAGACTGCTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGGTACTGCTGATAGG 300

QY 301 GTGCTTCGAGTGCCTCCCGGAGGTCTCTGATACCGTGCAACATGAGCAAGATCCATAAC 360  
 DB 301 GTGCTTCGAGTGCCTCCCGGAGGTCTCTGATACCGTGCAACATGAGCAAGATCCATAAC 360

QY 361 CTCAGAGAAACCAAGAGGCGGCCATGATTGAACAGATGGATTGCACGAGTTCTC 420  
 DB 361 CTCAGAGAAACCAAGAGGCGGCCATGATTGAACAGATGGATTGCACGAGTTCTC 420

QY 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGATCGGCAACAACAGCAATTCGGTCT 480  
 DB 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGATCGGCAACAACAGCAATTCGGTCT 480

QY 481 CTGATGCGCGCTGTTCCGGTGTGAGCGAGGGCGGCCGGTCTTTTGTCAAGACCG 540  
 DB 481 CTGATGCGCGCTGTTCCGGTGTGAGCGAGGGCGGCCGGTCTTTTGTCAAGACCG 540

QY 541 ACCTGTCCGTCCTTGAATGAATGACCTGACGAGGAGGCGGCTATCGTGGCTGGCCA 600  
 DB 541 ACCTGTCCGTCCTTGAATGAATGACCTGACGAGGAGGCGGCTATCGTGGCTGGCCA 600

QY 601 CGACGGGCTTCTTGGCGAGCTGTGCTCGAGTGTGTCACTGAAAGCGGAAAGGACTGCG 660  
 DB 601 CGACGGGCTTCTTGGCGAGCTGTGCTCGAGTGTGTCACTGAAAGCGGAAAGGACTGCG 660

QY 661 TGCTATTGGGCGAAGTGCCTGGGGCAGGATCTCTGTCTCATCTCACCTTCTCTCCCGAGA 720  
 DB 661 TGCTATTGGGCGAAGTGCCTGGGGCAGGATCTCTCTGTCTCATCTCACCTTCTCTCCCGAGA 720

QY 721 AAGTATCCATCATGGCTGATGCAATGGGGCTGCATACGCTTGTGATCCGGCTACTCTGCC 780  
 DB 721 AAGTATCCATCATGGCTGATGCAATGGGGCTGCATACGCTTGTGATCCGGCTACTCTGCC 780

QY 781 CATTCGACCAACCAAGCGAATCATCGATCGAGCGAGCACGTACTCGGATGAAAGCGGTC 840  
 DB 781 CATTCGACCAACCAAGCGAATCATCGATCGAGCGAGCACGTACTCGGATGAAAGCGGTC 840

QY 841 TTGTCGATCAGGATGATCTGCAAGAGCATAGGGGCTCGCGCAGCGAAGCTGTCG 900  
 DB 841 TTGTCGATCAGGATGATCTGCAAGAGCATAGGGGCTCGCGCAGCGAAGCTGTCG 900

QY 901 CAGGCTCAAGGGCGGATGCGCCGACGAGGATCTCGTGACCCATCGCGATGCTG 960  
 DB 901 CAGGCTCAAGGGCGGATGCGCCGACGAGGATCTCGTGACCCATCGCGATGCTG 960

QY 961 GCTTCCGGAATATCATGCTGGAAATGCGCGCTTTTCTGGATTCATCGACTTGGCCGCG 1020  
 DB 961 GCTTCCGGAATATCATGCTGGAAATGCGCGCTTTTCTGGATTCATCGACTTGGCCGCG 1020

QY 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATTTGCTGAAGAGC 1080  
 DB 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATTTGCTGAAGAGC 1080

QY 1081 TTGGCGGGAATGGGCTGACCGCTTCTCGTCTTACCGTATCGCGCTCCCGATTCG 1140  
 DB 1081 TTGGCGGGAATGGGCTGACCGCTTCTCGTCTTACCGTATCGCGCTCCCGATTCG 1140

QY 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTTCCTTGAGTTTAAACAGACCAACG 1200  
 DB 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTTCCTTGAGTTTAAACAGACCAACG 1200

QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCTAACGTTACTGCG 1260  
 DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCTAACGTTACTGCG 1260

QY 1261 CGAAGCGCTTGAATAAGCGCGTGTGGTGTCTATATGTTATTTTCCACCATATTG 1320  
 DB 1261 CGAAGCGCTTGAATAAGCGCGTGTGGTGTCTATATGTTATTTTCCACCATATTG 1320

QY 1321 CCGTCTTTTGGCAATGTAGGGCGCGGAAACCTTGCCCTCTCTTCTTGAAGAGTTCCT 1380  
 DB 1321 CCGTCTTTTGGCAATGTAGGGCGCGGAAACCTTGCCCTCTCTTCTTGAAGAGTTCCT 1380

QY 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAAATGTCGTAAGAGCA 1440  
 DB 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGAAATGTCGTAAGAGCA 1440

QY 1441 GTTCTCTGGAAGTCTTCTGAAGCAACAAACGTCGTAGCGACCTTTTGACGAGCGG 1500  
 DB 1441 GTTCTCTGGAAGTCTTCTGAAGCAACAAACGTCGTAGCGACCTTTTGACGAGCGG 1500

QY 1501 AACCCCGACCTGCGGACAGGTGCTCTGCGGCGCAAAAGCAACGTCGTATAGATACACT 1560  
 DB 1501 AACCCCGACCTGCGGACAGGTGCTCTGCGGCGCAAAAGCAACGTCGTATAGATACACT 1560

QY 1561 GCAAGAGCGGCAACCCCGAGTGCACGTTGTGAGTTGGAAGTGTGGAAGAGTCAAA 1620  
 DB 1561 GCAAGAGCGGCAACCCCGAGTGCACGTTGTGAGTTGGAAGTGTGGAAGAGTCAAA 1620

QY 1621 TGGCTCTCTCAAGCGTATTCAAAGGGCTGAAGGATGCCAGAGGTATCCCATTTG 1680  
 DB 1621 TGGCTCTCTCAAGCGTATTCAAAGGGCTGAAGGATGCCAGAGGTATCCCATTTG 1680

QY 1681 ATGGATCTGATCTGGGGCTCGGTGCAATGCTTTTACATGTTTGTAGTTCGAGGTTAAA 1740  
 DB 1681 ATGGATCTGATCTGGGGCTCGGTGCAATGCTTTTACATGTTTGTAGTTCGAGGTTAAA 1740

QY 1741 AACGTCTAGGCCCCCGGAAACCGAGGGAGCGTGGTTTCTTTTGAAACACGATAATACC 1800

Db 1741 AACGTCAGGCCCCCGAACCAACGGGACGTTGGTTTCTTTTCAAAAAACAGATATACC 1800  
QY 1801 ATGGGCGCTATTACGGCCTACTCCAAACAGACGCGAGGCTACTTGGCTGCATCATCACT 1860  
Db 1801 ATGGGCGCTATTACGGCCTACTCCAAACAGACGCGAGGCTACTTGGCTGCATCATCACT 1860  
QY 1861 AGCCTCACAGGCGGGACAGGAACCAAGGTCGAGGGGAGGTCAAGTGTCTCCACCGCA 1920  
Db 1861 AGCCTCACAGGCGGGACAGGAACCAAGGTCGAGGGGAGGTCAAGTGTCTCCACCGCA 1920  
QY 1921 ACACAATCTTCTCGGACCTGCGTCAATGCGGTGTGTGACTGTCTATCATGTTGCC 1980  
Db 1921 ACACAATCTTCTCGGACCTGCGTCAATGCGGTGTGTGACTGTCTATCATGTTGCC 1980  
QY 1981 GGCTCAAAAGACCTTTCGCGGCCCAAGGCCCAATCACCCAAATGTACACCAATGTGAC 2040  
Db 1981 GGCTCAAAAGACCTTTCGCGGCCCAAGGCCCAATCACCCAAATGTACACCAATGTGAC 2040  
QY 2041 CAGGACCTCGTGGGTGCAAGCGCCCCCGGGGCGGTCTCTTGACACCAATGACACCTGC 2100  
Db 2041 CAGGACCTCGTGGGTGCAAGCGCCCCCGGGGCGGTCTCTTGACACCAATGACACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTGGTACAGAGGATGCGGATGCAATTCGGGTGCGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTGGTACAGAGGATGCGGATGCAATTCGGGTGCGCGCGG 2160  
QY 2161 GGCACACAGCGGGGAGGCTACTCTCCCGCCAGGCGCGTCTCTACTTGAAGGGCTCTTCG 2220  
Db 2161 GGCACACAGCGGGGAGGCTACTCTCCCGCCAGGCGCGTCTCTACTTGAAGGGCTCTTCG 2220  
QY 2221 GGGGCTGCACTGCTGCGGTCGAGGCGACGCTGTGGGCACTTTTGGGCTGCGGTGTC 2280  
Db 2221 GGGGCTGCACTGCTGCGGTCGAGGCGACGCTGTGGGCACTTTTGGGCTGCGGTGTC 2280  
QY 2281 ACCGAGGGGTGCAAGGGGTGGAATTGTTACCGTTCGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCGAGGGGTGCAAGGGGTGGAATTGTTACCGTTCGAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGCTTTCACGGCAACTCGTCCCTCCCGCGGTACGAGTCCGGTATGCA 2460  
Db 2341 CGGTCCCGGCTTTCACGGCAACTCGTCCCTCCCGCGGTACGAGTCCGGTATGCA 2460  
QY 2401 GCCCATCTACACGCCCTACTGTAGCGGCAAGGACACTAAGTTCGGTATGCA 2460  
Db 2401 GCCCATCTACACGCCCTACTGTAGCGGCAAGGACACTAAGTTCGGTATGCA 2460  
QY 2461 GCCCAAGGGTATAAGGTGCTTGTCTGAAACCGGTGCGTCCCGCCACCTTAGGTTTCGGG 2520  
Db 2461 GCCCAAGGGTATAAGGTGCTTGTCTGAAACCGGTGCGTCCCGCCACCTTAGGTTTCGGG 2520  
QY 2521 GGGTATATGCTTAAGGCACATGTTATCGACCTAATGGAAGTTTCTTCCCGACGTTGTC 2580  
Db 2521 GGGTATATGCTTAAGGCACATGTTATCGACCTAATGGAAGTTTCTTCCCGACGTTGTC 2580  
QY 2581 ACACGGGTGCCCCCATACGTAATGATGAGTGCACCTCAACTGACTCGACCACT 2700  
Db 2581 ACACGGGTGCCCCCATACGTAATGATGAGTGCACCTCAACTGACTCGACCACT 2700  
QY 2641 TCTGGGGCGCTATGACATCATATATGATGATGAGTGCACCTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGCGCTATGACATCATATATGATGATGAGTGCACCTCAACTGACTCGACCACT 2700  
QY 2701 ATCTTGGGCGCTATGACATCATATATGATGATGAGTGCACCTCAACTGACTCGACCACT 2760  
Db 2701 ATCTTGGGCGCTATGACATCATATATGATGATGAGTGCACCTCAACTGACTCGACCACT 2760  
QY 2761 CTCGCCACCGCTACGCTCCGGGATCGGTACCGTGCACATCCAAACATCCAGAGGTG 2820  
Db 2761 CTCGCCACCGCTACGCTCCGGGATCGGTACCGTGCACATCCAAACATCCAGAGGTG 2820  
QY 2821 GCTCTGTGAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCAATC 2880  
Db 2821 GCTCTGTGAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCAATC 2880

Db 2821 GCTCTGTGAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCAATC 2880  
QY 2881 AAGGGGGGAGGACACCTCATTTTCTGCCATTCCAAGAAAGAAATGTGATGAGTCCGGCG 2940  
Db 2881 AAGGGGGGAGGACACCTCATTTTCTGCCATTCCAAGAAAGAAATGTGATGAGTCCGGCG 2940  
QY 2941 AAGCTGTCCGGCTCGACTCAATGCTGAGCATATTAACGGGGCTTGTGATGATCCGTC 3000  
Db 2941 AAGCTGTCCGGCTCGACTCAATGCTGAGCATATTAACGGGGCTTGTGATGATCCGTC 3000  
QY 3001 ATACCAACTAGCGGAGAGCTCATTTGCTGAGCAACGCGCTCTAATGACGGCTTTTACC 3060  
Db 3001 ATACCAACTAGCGGAGAGCTCATTTGCTGAGCAACGCGCTCTAATGACGGCTTTTACC 3060  
QY 3061 GGCAGTTTTCGACTCAGTGCATGCACTGCAATCATGTGTCACCCAGACAGTGCAGTTCAGC 3120  
Db 3061 GGCAGTTTTCGACTCAGTGCATGCACTGCAATCATGTGTCACCCAGACAGTGCAGTTCAGC 3120  
QY 3121 CTGACCCGACCTTCAACCAATTGAGACGACGCGGTGCCAAGACGCGGTGTCACGCTCG 3180  
Db 3121 CTGACCCGACCTTCAACCAATTGAGACGACGCGGTGCCAAGACGCGGTGTCACGCTCG 3180  
QY 3181 CAGCGGCGAGGACGAGCTGTGAGGGGAGGATGGGCATTTACAGTTTGTGACTCCAGGA 3240  
Db 3181 CAGCGGCGAGGACGAGCTGTGAGGGGAGGATGGGCATTTACAGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGGGCTCGGGCATGTTCCGATTCTGTCGAGTGTCTGTCGAGTGTCTTACAGGGCTC 3300  
Db 3241 GAACGGGCTCGGGCATGTTCCGATTCTGTCGAGTGTCTTACAGGGCTC 3300  
QY 3301 GCTTGTGACGAGCTCACGCGCGCGAGACCTCACTAGTTAGTTGCGGGCTTACTTAAACACA 3360  
Db 3301 GCTTGTGACGAGCTCACGCGCGCGAGACCTCACTAGTTAGTTGCGGGCTTACTTAAACACA 3360  
QY 3361 CCAGGGTTGCGGCTCTGCCAGGACCATCTGAGATTCTGGAGAGCGTCTTTACAGGGCTC 3420  
Db 3361 CCAGGGTTGCGGCTCTGCCAGGACCATCTGAGATTCTGGAGAGCGTCTTTACAGGGCTC 3420  
QY 3421 ACCCATAGACGCCCATTTCTTCCAGACTAAGCAGGCGAGGACAGTTCCTTACAGGGCTC 3480  
Db 3421 ACCCATAGACGCCCATTTCTTCCAGACTAAGCAGGCGAGGACAGTTCCTTACAGGGCTC 3480  
QY 3481 CTGGTACGATACAGGCTACGCTGTCGCGAGGGCTCAGGCTCCACCTCATCGTGGGAC 3540  
Db 3481 CTGGTACGATACAGGCTACGCTGTCGCGAGGGCTCAGGCTCCACCTCATCGTGGGAC 3540  
QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAAGGGCCAAACGCCCTGCTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCAAGGGCCAAACGCCCTGCTG 3600  
QY 3601 TATAGGCTGGGAGCGCTTCAAAACGAGGTTACTACCCACACCCCATACCAATAATCATC 3660  
Db 3601 TATAGGCTGGGAGCGCTTCAAAACGAGGTTACTACCCACACCCCATACCAATAATCATC 3660  
QY 3661 ATGGCATGATCGGCTGACCTGAGGTCGTCACGAGCACTGGTGTGTTAGGCGGA 3720  
Db 3661 ATGGCATGATCGGCTGACCTGAGGTCGTCACGAGCACTGGTGTGTTAGGCGGA 3720  
QY 3721 GTCTTAGCAGCTCTGGCGCGTATTTGCCCTGACCAAGCAGGCGTGTCTTGTGGGCGAG 3780  
Db 3721 GTCTTAGCAGCTCTGGCGCGTATTTGCCCTGACCAAGCAGGCGTGTCTTGTGGGCGAG 3780  
QY 3781 ATCATCTTGTTCGGAAACCGGCCATCATTCGAGACGGGAAGTCTTTTACGGGAGTTC 3840  
Db 3781 ATCATCTTGTTCGGAAACCGGCCATCATTCGAGACGGGAAGTCTTTTACGGGAGTTC 3840  
QY 3841 GATGAGTGGAGAGTGGCGCTCACACCTCCTTACATCGAAGAGGAATGAGTTCGCTC 3900  
Db 3841 GATGAGTGGAGAGTGGCGCTCACACCTCCTTACATCGAAGAGGAATGAGTTCGCTC 3900  
QY 3901 GAAACAATTCAAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAGGAGGCT 3960  
Db 3901 GAAACAATTCAAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAGGAGGCT 3960

QY	3961	GCTGCTCCCGTGGTGGAAATCCCAAGTGGCGGACCCCTCGAAGCCCTCTGGGCGAAGCATATG	4020
Db	3961	GCTGCTCCCGTGGTGGAAATCCCAAGTGGCGGACCCCTCGAAGCCCTCTGGGCGAAGCATATG	4020
QY	4021	TGGAAATTTCACTACGCGGGATACAAATATTTAGCAGGCTTTGCCACTCTGCTGGCAACCCC	4080
Db	4021	TGGAAATTTCACTACGCGGGATACAAATATTTAGCAGGCTTTGCCACTCTGCTGGCAACCCC	4080
QY	4081	GCGATAGCATCACTGATGGGATTCACAGCTCTATCACAGAGCCGCTCAACCCCAACAT	4140
Db	4081	GCGATAGCATCACTGATGGGATTCACAGCTCTATCACAGAGCCGCTCAACCCCAACAT	4140
QY	4141	ACCCCTCGTGTAAACATCCTGGGGGATGGTGGCGGCCCAACTTGCTCTCTCCACAGCGCT	4200
Db	4141	ACCCCTCGTGTAAACATCCTGGGGGATGGTGGCGGCCCAACTTGCTCTCTCCACAGCGCT	4200
QY	4201	GCTTCTGCTTTCTGTPAGCGCGCGCATCGCTGGAGCGGCTTTTGGCAGCATAGGCCCTTGGG	4260
Db	4201	GCTTCTGCTTTCTGTPAGCGCGCGCATCGCTGGAGCGGCTTTTGGCAGCATAGGCCCTTGGG	4260
QY	4261	AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGCGCGCTGCTGGCC	4320
Db	4261	AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGCGCGCTGCTGGCC	4320
QY	4321	TTTAAAGTTCATAGCGCGAGATGCCCTCCACAGGAGACCTGTGTTAACTACTCCCTGCT	4380
Db	4321	TTTAAAGTTCATAGCGCGAGATGCCCTCCACAGGAGACCTGTGTTAACTACTCCCTGCT	4380
QY	4381	ATCCTCTCCCTGGCGCCCTAGTCTGGGGTGTGTGCGCAGCGATCTGCTGCGGCAC	4440
Db	4381	ATCCTCTCCCTGGCGCCCTAGTCTGGGGTGTGTGCGCAGCGATCTGCTGCGGCAC	4440
QY	4441	GTGGGCCACAGGGAGCGGGCTGTGCAGTGGATGAGACCGGCTGATAGCGTTGCGTTGCGG	4500
Db	4441	GTGGGCCACAGGGAGCGGGCTGTGCAGTGGATGAGACCGGCTGATAGCGTTGCGTTGCGG	4500
QY	4501	GGTAACCAAGTCTCCCCACGACATATGCTGCTGAGAGCGACGCTGACGACGTTGTCAC	4560
Db	4501	GGTAACCAAGTCTCCCCACGACATATGCTGCTGAGAGCGACGCTGACGACGTTGTCAC	4560
QY	4561	CAGATCCTCTCTAGTCTTACCATCACTCAGTCTGTGTAAGAGGCTTACCAGTGTATCAAC	4620
Db	4561	CAGATCCTCTCTAGTCTTACCATCACTCAGTCTGTGTAAGAGGCTTACCAGTGTATCAAC	4620
QY	4621	GAGGACTGTCCACGCCATGCTCCGGCTGTGGCTTAAGAGATGTTTGGATTTGGATATGC	4680
Db	4621	GAGGACTGTCCACGCCATGCTCCGGCTGTGGCTTAAGAGATGTTTGGATTTGGATATGC	4680
QY	4681	ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCTCCCGGATTCGCGGA	4740
Db	4681	ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCTCCCGGATTCGCGGA	4740
QY	4741	GTCCCTCTTCTCATGTCAAGCTGGCTCAAGGAGTCTGCGGGGCGACGGCATCATG	4800
Db	4741	GTCCCTCTTCTCATGTCAAGCTGGCTCAAGGAGTCTGCGGGGCGACGGCATCATG	4800
QY	4801	CAAAACCACTGCCCATGTGGAGCACAGATCACCGGACATGTGAATAACGGTTCCATGAGG	4860
Db	4801	CAAAACCACTGCCCATGTGGAGCACAGATCACCGGACATGTGAATAACGGTTCCATGAGG	4860
QY	4861	ATCGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAAACATTCGCCATTTAAACGGGTAC	4920
Db	4861	ATCGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAAACATTCGCCATTTAAACGGGTAC	4920
QY	4921	ACACAGGGCCCTGACAGCCCTCCCGGCGCCTAAATTTCTAGGGGCTGTGGCGGGTG	4980
Db	4921	ACACAGGGCCCTGACAGCCCTCCCGGCGCCTAAATTTCTAGGGGCTGTGGCGGGTG	4980
QY	4981	GCTCTGAGGAGTACGTGGAGGTTACCGGGTGGGGGATTTCCACTACGTGACGGGCATG	5040
Db	4981	GCTCTGAGGAGTACGTGGAGGTTACCGGGTGGGGGATTTCCACTACGTGACGGGCATG	5040
QY	5041	ACCACTGACAAGTAAAGTGCCTGTGACGTTCCGCGCCCGCCGAAATTTCTTCAAGAAGTG	5100
Db	5041	ACCACTGACAAGTAAAGTGCCTGTGACGTTCCGCGCCCGCCGAAATTTCTTCAAGAAGTG	5100
QY	5101	GATGGGTGGCGTTTGCAACAGGTACGCTCCAGCGTGCAGAGCCCTCTTACCGGAGGAGGTC	5160
Db	5101	GATGGGTGGCGTTTGCAACAGGTACGCTCCAGCGTGCAGAGCCCTCTTACCGGAGGAGGTC	5160
QY	5161	ACATTTCTGTGTCGGGCTCAATCAATAGCTGTTGGGTACAGCTCCCATCGAGCCCGAA	5220
Db	5161	ACATTTCTGTGTCGGGCTCAATCAATAGCTGTTGGGTACAGCTCCCATCGAGCCCGAA	5220
QY	5221	CGGACGTAGCAGTGTCTCCTTCCATGCTCACCGACCCCTCCACATTTACGCGGAGACG	5280
Db	5221	CGGACGTAGCAGTGTCTCCTTCCATGCTCACCGACCCCTCCACATTTACGCGGAGACG	5280
QY	5281	GCTAAGCGTAGGCTGGCCAGGGATCTCCCGCTCTCTTGGCAGCTCATCAGCTAGCCAG	5340
Db	5281	GCTAAGCGTAGGCTGGCCAGGGATCTCCCGCTCTCTTGGCAGCTCATCAGCTAGCCAG	5340
QY	5341	CTGTCTGGCGCTTCTTGAAGGCAACATGCACTACCCGTATGACTCCCCGACGCTGAC	5400
Db	5341	CTGTCTGGCGCTTCTTGAAGGCAACATGCACTACCCGTATGACTCCCCGACGCTGAC	5400
QY	5401	CTCATCGAGGCCAACTCTGTGGCGGAGAGATGGCGGGGAAATCACCCGCTGGAG	5460
Db	5401	CTCATCGAGGCCAACTCTGTGGCGGAGAGATGGCGGGGAAATCACCCGCTGGAG	5460
QY	5461	TGAGAAATTAAGGTAGTAAATTTGGACTCTTTTGGAGCCGCTTCAAGCGGAGGAGATGAG	5520
Db	5461	TGAGAAATTAAGGTAGTAAATTTGGACTCTTTTGGAGCCGCTTCAAGCGGAGGAGATGAG	5520
QY	5521	AGGAAAGTATCGTTCGCGCGGAGATCTGCGGAGGTCCAGGAAATTTCCCTCGACCGATG	5580
Db	5521	AGGAAAGTATCGTTCGCGCGGAGATCTGCGGAGGTCCAGGAAATTTCCCTCGACCGATG	5580
QY	5581	CCCATATGGGCAACCGCGGATTAACAACCTTCCACTGTTAGAGTCTTGGAAAGACCCGAC	5640
Db	5581	CCCATATGGGCAACCGCGGATTAACAACCTTCCACTGTTAGAGTCTTGGAAAGACCCGAC	5640
QY	5641	TACGTCCCTCCAGTGTTACACGGGTGTCCATTGCCGCTTCCAGAGCCCTCCGATACCA	5700
Db	5641	TACGTCCCTCCAGTGTTACACGGGTGTCCATTGCCGCTTCCAGAGCCCTCCGATACCA	5700
QY	5701	CCTTCCAGGAGGAGGAGGTTGCTGTGAGATCTACCGTGTCTTCTGCTTGGCG	5760
Db	5701	CCTTCCAGGAGGAGGAGGTTGCTGTGAGATCTACCGTGTCTTCTGCTTGGCG	5760
QY	5761	GAGCTCGCCACAAAGACCTTTCGCGAGCTCCGAATCTGCGGCGTCCAGAGCGGCA	5820
Db	5761	GAGCTCGCCACAAAGACCTTTCGCGAGCTCCGAATCTGCGGCGTCCAGAGCGGCA	5820
QY	5821	ACGCGCTCTCTGACACAGCCCTCCGACGAGCGGAGTCCGAGCTTGGAGTCTGAT	5880
Db			



```

OS Synthetic.
XX Key location/Qualifiers
FH 5'UTR 1..341
FT /*tag= a
FT 342..1181
FT /*tag= b
FT /*product= "Core-neo fusion protein"
FT misc_signal 1190..1800
FT /*tag= c
FT /*label= IRES
FT /*note= "Internal ribosome entry site from EMCV"
FT CDS 1801..7758
FT /*tag= d
FT /*product= "Polyprotein"
FT /*note= "Comprising NS3, NS4A, NS4B, NS5A and NS5B"
FT mutation replace(4847,A)
FT /*tag= e
FT 7759..7991
FT /*tag= f
FT misc_feature 7992..10690
FT /*tag= g
FT /*note= "Plasmid derived sequences"
XX
XX WO200259321-A2.
XX
XX 01-AUG-2002.
XX
XX 16-JAN-2002; 2002WO-BP000526.
XX
XX 23-JAN-2001; 2001US-0263479P.
XX
XX (RICE-) IST RICERCHIE BIOL MOLECOLARE ANGELETTI.
XX
XX De Francesco R, Migliaccio G, Paonessa G;
XX
XX WPI; 2002-599793/64.
XX
XX New Hepatitis C virus (HCV) replicons comprising altered HCV NS3 or HCV
XX NS5 encoding region, or encephalomyocarditis virus (EMCV) internal
XX ribosome entry site (IRES) region, useful in studying HCV replication and
XX expression.
XX
XX Example 1; Page; 69pp; English.
XX
XX The invention relates to nucleic acid molecules comprising altered HCV
XX NS3 or HCV NS5 encoding region, or encephalomyocarditis virus (EMCV)
XX internal ribosome entry site (IRES) region coding for one or more NS3,
XX NS5A, or EMCV IRES mutations, respectively. The location of the mutations
XX are detailed in the specification. Also included are (1) an expression
XX vector comprising a nucleotide sequence coding for the altered nucleic
XX acids, which is transcriptionally coupled to an exogenous promoter; (2) a
XX recombinant cell human hepatoma cell comprising the altered nucleic acids
XX ; (3) a recombinant cell produced by introducing into a human hepatoma
XX cell the altered nucleic acids; (4) producing an HCV (hepatitis C virus)
XX replicon enhanced cell or which containing a functional HCV replicon; (5)
XX an HCV replicon enhanced cells made in the method; and (6) measuring the
XX ability of a compound to affect HCV activity. The HCV replicons and HCV
XX replicon enhanced cells are useful in studying HCV replication and
XX expression, and HCV and host cell interactions, producing HCV RNA and
XX proteins, and providing a system for measuring the ability of a compound
XX to modulate one or more HCV activities e.g. to discover drugs which may
XX treat HCV mediated diseases such as liver failure, cirrhosis and
XX hepatocellular carcinoma. The present sequence is an HCV based vector
XX pHCVNeo.17 mutant of the invention. Note: The present sequence is not
XX shown in the specification but was created by the indexer using the HCV
XX vector sequence appearing as ABK91412 and the information in example 1
XX
XX Sequence 10690 BP; 2333 A; 3046 C; 2908 G; 2403 T; 0 U; 0 Other;
XX
XX Query Match 100.0%; Score 7985.8; DB 6; Length 10690;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 7987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```





Db 3241 GAAAGGCGCTCGGAGATGTTGATTCCTCGGTTCTGTGGAGTGCTATGACGCGGGCTGT 3300  
 QY 3301 GCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
 Db 3301 GCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
 QY 3361 CGAGGGTGGCGCTGTCCAGAGACCATCTGGAGTTCTGGGAGAGGCTTTTACAGGCTC 3420  
 Db 3361 CGAGGGTGGCGCTGTCCAGAGACCATCTGGAGTTCTGGGAGAGGCTTTTACAGGCTC 3420  
 QY 3421 ACCACATAGAGCGCCATTCTTGTGCCAGATAAGACGAGCAGAGACAACCTTCCCTAC 3480  
 Db 3421 ACCACATAGAGCGCCATTCTTGTGCCAGATAAGACGAGCAGAGACAACCTTCCCTAC 3480  
 QY 3481 CTGTAGCATACAGGCTACGTTGCGGCGAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
 Db 3481 CTGTAGCATACAGGCTACGTTGCGGCGAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
 QY 3541 CAAATGTGAAGTGTCTATACGGCTAAAGCCTACGCTGACGGGCCAACGCCCTGCTG 3600  
 Db 3541 CAAATGTGAAGTGTCTATACGGCTAAAGCCTACGCTGACGGGCCAACGCCCTGCTG 3600  
 QY 3601 TATAGGCTGGAGCGCTTCAAACAGAGGTTACTACACACACCCCAATAACCAATACATC 3660  
 Db 3601 TATAGGCTGGAGCGCTTCAAACAGAGGTTACTACACACACCCCAATAACCAATACATC 3660  
 QY 3661 ATGGCATGATGCTGGCTGACCTGGAGGTCGTACAGACACCTGGGTGCTGTAGCGGA 3720  
 Db 3661 ATGGCATGATGCTGGCTGACCTGGAGGTCGTACAGACACCTGGGTGCTGTAGCGGA 3720  
 QY 3721 GTCTAGAGCTGTGCGCGGTTATGCTGACAAACAGGAGCGTGTCTATTTGGGCGAG 3780  
 Db 3721 GTCTAGAGCTGTGCGCGGTTATGCTGACAAACAGGAGCGTGTCTATTTGGGCGAG 3780  
 QY 3781 ATCATCTTGTCCGGAAGCGGCGATCAITCCCGACAGGGAAGTCTTTACCGGGAGTTC 3840  
 Db 3781 ATCATCTTGTCCGGAAGCGGCGATCAITCCCGACAGGGAAGTCTTTACCGGGAGTTC 3840  
 QY 3841 GATGAGATGAAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCGAGCTGCC 3900  
 Db 3841 GATGAGATGAAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGCGAGCTGCC 3900  
 QY 3901 GAACAATCAAACAGAGGAATCGGGTCTGCAACAGCCCAAGCAAGCGGAGGCT 3960  
 Db 3901 GAACAATCAAACAGAGGAATCGGGTCTGCAACAGCCCAAGCAAGCGGAGGCT 3960  
 QY 3961 GCTGCTCCGTTGGTGAATCCAGTGGCGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
 Db 3961 GCTGCTCCGTTGGTGAATCCAGTGGCGGACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
 QY 4021 TGGAAATTCATAGCGGGAATCAATATTTAGAGGCTTTGCTACTCTGCTGGCAACCCC 4080  
 Db 4021 TGGAAATTCATAGCGGGAATCAATATTTAGAGGCTTTGCTACTCTGCTGGCAACCCC 4080  
 QY 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTTCTATCACAGCCCGCTCAACCCCAACAT 4140  
 Db 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTTCTATCACAGCCCGCTCAACCCCAACAT 4140  
 QY 4141 ACCCTCTGTTTAACTCTCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
 Db 4141 ACCCTCTGTTTAACTCTCTGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
 QY 4201 GCTTCTGTTTCTAGAGCGCGGATCGCTGGAGCGGCTGTGSCAGCATAGGCTTGGG 4260  
 Db 4201 GCTTCTGTTTCTAGAGCGCGGATCGCTGGAGCGGCTGTGSCAGCATAGGCTTGGG 4260  
 QY 4261 AAGGTGCTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGGCTGTGGCC 4320  
 Db 4261 AAGGTGCTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGGCTGTGGCC 4320  
 QY 4321 TTTAAGGTGATGAGCGGCGAGATGCTCCACCGAGGACCTGGTTAAGCTACTCCCTGCT 4380  
 Db 4321 TTTAAGGTGATGAGCGGCGAGATGCTCCACCGAGGACCTGGTTAAGCTACTCCCTGCT 4380

QY 4381 ATCTCTCCCTGCGCGCTAGTCTCGGGGTCTGTGTCGACGACATACTGCGTGGGCAC 4440  
 Db 4381 ATCTCTCCCTGCGCGCTAGTCTCGGGGTCTGTGTCGACGACATACTGCGTGGGCAC 4440  
 QY 4441 GTGGGCCAGGGAGGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTCCGTTCCGG 4500  
 Db 4441 GTGGGCCAGGGAGGGGGCTGTGAGTGGATGAACCGGCTGATAGCGTTCCGTTCCGG 4500  
 QY 4501 GGTACACAGCTCTCCCGACGACATATGTGCTGAGAGCAGCGCTGACAGCGTGTCACT 4560  
 Db 4501 GGTACACAGCTCTCCCGACGACATATGTGCTGAGAGCAGCGCTGACAGCGTGTCACT 4560  
 QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACGAGTGGATCAAC 4620  
 Db 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACGAGTGGATCAAC 4620  
 QY 4621 GAGGACTGTCCAGGCACTGCTCGGCTGCTGAAGAGATGTTTGGGATGGATATGC 4680  
 Db 4621 GAGGACTGTCCAGGCACTGCTCGGCTGCTGAAGAGATGTTTGGGATGGATATGC 4680  
 QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCGCGGATTTGCCGGA 4740  
 Db 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCTGCGCGGATTTGCCGGA 4740  
 QY 4741 GTCCCTCTCTCATGTCAACGTGGGTAAAGGAGTCTGGCGGGGCGACGCGCATCATG 4800  
 Db 4741 GTCCCTCTCTCATGTCAACGTGGGTAAAGGAGTCTGGCGGGGCGACGCGCATCATG 4800  
 QY 4801 CAAACACCTGCGGATGAGGACCTGTAGTAAACGTGCAATGGAACGTTCCATGAGG 4860  
 Db 4801 CAAACACCTGCGGATGAGGACCTGTAGTAAACGTGCAATGGAACGTTCCATGAGG 4860  
 QY 4861 ATCGTGGGCGCTAGGACCTGTAGTAAACGTGCAATGGAACGTTCCATGAGG 4920  
 Db 4861 ATCGTGGGCGCTAGGACCTGTAGTAAACGTGCAATGGAACGTTCCATGAGG 4920  
 QY 4921 ACCACGCGCGCTGACGCGCTTCCCGGCGCCAAATTTATTTAGGGGCTGTGGCGGCTG 4980  
 Db 4921 ACCACGCGCGCTGACGCGCTTCCCGGCGCCAAATTTATTTAGGGGCTGTGGCGGCTG 4980  
 QY 4981 GCTGCTGAGAGTACGTGGAGGTTAGCGGGTGGGGATTTCCACTACGTGACGGGCGATG 5040  
 Db 4981 GCTGCTGAGAGTACGTGGAGGTTAGCGGGTGGGGATTTCCACTACGTGACGGGCGATG 5040  
 QY 5041 ACCACTGACAACTGAAGTCCCGTGTACAGTTCCGGCCCGGAAATTTCTTACAGAAAGTG 5100  
 Db 5041 ACCACTGACAACTGAAGTCCCGTGTACAGTTCCGGCCCGGAAATTTCTTACAGAAAGTG 5100  
 QY 5101 GATGGGCTGCGGTTGACAGGTTACGCTCCAGCGTGAACCCCTCTCTACGGGAGGAGTTC 5160  
 Db 5101 GATGGGCTGCGGTTGACAGGTTACGCTCCAGCGTGAACCCCTCTCTACGGGAGGAGTTC 5160  
 QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGGTGGGTACAGCTCCCATGCGAGCCGAA 5220  
 Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGGTGGGTACAGCTCCCATGCGAGCCGAA 5220  
 QY 5221 CCGGACGTAGCAGTGTCACTTCCATGCTCAACGACCCCTCCCACTATAGCGGCGAGACG 5280  
 Db 5221 CCGGACGTAGCAGTGTCACTTCCATGCTCAACGACCCCTCCCACTATAGCGGCGAGACG 5280  
 QY 5281 GCTAAGGCTAGGCTGGCGAGGGATCTCCCGCTCTTGGCCAGCTCATGAGTGGCAG 5340  
 Db 5281 GCTAAGGCTAGGCTGGCGAGGGATCTCCCGCTCTTGGCCAGCTCATGAGTGGCAG 5340  
 QY 5341 CTGCTGGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGCGTGCAC 5400  
 Db 5341 CTGCTGGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGCGTGCAC 5400  
 QY 5401 CTGCTGGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGCGTGCAC 5460  
 Db 5401 CTGCTGGCGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGCGTGCAC 5460





QY 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACACAGACAATTCGGCTGCT 480  
DB 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACACAGACAATTCGGCTGCT 480  
QY 481 CTGATGCGCGCTGTTCGGCTGTACGCGAGGGGCGCCCGGCTCTTTTGTGCAAGACCG 540  
DB 481 CTGATGCGCGCTGTTCGGCTGTACGCGAGGGGCGCCCGGCTCTTTTGTGCAAGACCG 540  
QY 541 ACCTGTCCGGTCCCTGAACTGAATGCACTGACGAGGAGGAGGAGCGGCTATTCGTGGTGGCCA 600  
DB 541 ACCTGTCCGGTCCCTGAACTGAATGCACTGACGAGGAGGAGGAGCGGCTATTCGTGGTGGCCA 600  
QY 601 CGACGGGCTTCCTTGGCGACGTGTCTGCTCAGCTTGTCTACCTGAAGCGGAGGAGGAGTGGC 660  
DB 601 CGACGGGCTTCCTTGGCGACGTGTCTGCTCAGCTTGTCTACCTGAAGCGGAGGAGGAGTGGC 660  
QY 661 TGTCTATTGGGCGAAGTGGCGGGGAGGATCTCTGTCTATCTCACTTGTCTCTCTCCCGAGA 720  
DB 661 TGTCTATTGGGCGAAGTGGCGGGGAGGATCTCTGTCTATCTCACTTGTCTCTCTCCCGAGA 720  
QY 721 AAGTATCCATCATCGCTGATGCAATGCGGCGGTGCTATGCTTGTGATCGGCTACCTGTC 780  
DB 721 AAGTATCCATCATCGCTGATGCAATGCGGCGGTGCTATGCTTGTGATCGGCTACCTGTC 780  
QY 781 CATTCGACCAACAGCGAATCATCGCATCGAGCGAGCAGCTACTCGATGGAAGCGCGTC 840  
DB 781 CATTCGACCAACAGCGAATCATCGCATCGAGCGAGCAGCTACTCGATGGAAGCGCGTC 840  
QY 841 TTGTGCTATCAGATGATCTCGAAGAGCATCAGGGGCTCGGCCAGCCGAACTGTTCG 900  
DB 841 TTGTGCTATCAGATGATCTCGAAGAGCATCAGGGGCTCGGCCAGCCGAACTGTTCG 900  
QY 901 CCAGGCTCAAGGCGCGATCGCCGACGGCAGGATCTCTGCTGACCCATGGCGATGCCCT 960  
DB 901 CCAGGCTCAAGGCGCGATCGCCGACGGCAGGATCTCTGCTGACCCATGGCGATGCCCT 960  
QY 961 GCTTGGCGAATATCATGTTGGAATGGCGGCTTTCTGATTCATCGACTGTGGCGGC 1020  
DB 961 GCTTGGCGAATATCATGTTGGAATGGCGGCTTTCTGATTCATCGACTGTGGCGGC 1020  
QY 1021 TGGGTGCGGACCGGTATCAGACATAGCTTGGCTACCCGTGATTTGCTGAAGAGC 1080  
DB 1021 TGGGTGCGGACCGGTATCAGACATAGCTTGGCTACCCGTGATTTGCTGAAGAGC 1080  
QY 1081 TTGGCGCGAATGGGCTGACCGCTTCTCTGCTTTACGGTATCGCGCTCCCGATTCG 1140  
DB 1081 TTGGCGCGAATGGGCTGACCGCTTCTCTGCTTTACGGTATCGCGCTCCCGATTCG 1140  
QY 1141 AGCGATCGCTTCTATCGCTTCTTTGACGAGTTCTTCTGAGTTTAAACAGACCAACG 1200  
DB 1141 AGCGATCGCTTCTATCGCTTCTTTGACGAGTTCTTCTGAGTTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCTCCCTCCCGCCCTAAACGTTACTGGC 1260  
DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCTCCCTCCCGCCCTAAACGTTACTGGC 1260  
QY 1261 CGAAGCGCTTGAATAAGCGGCTGCTTCTATGATTTATTTTCCACCATTTG 1320  
DB 1261 CGAAGCGCTTGAATAAGCGGCTGCTTCTATGATTTATTTTCCACCATTTG 1320  
QY 1321 CGGTCTTTTGGCAATGTAGGGCCCGGAACTCTGCGCTCTCTCCCGCCCTAAACGTTACTGGC 1380  
DB 1321 CGGTCTTTTGGCAATGTAGGGCCCGGAACTCTGCGCTCTCTCTGACGAGCATTCCT 1380  
QY 1381 AGGGTCTTTTCCCTCTCGCCAAAGAAATGCAAGTCTGTTGAATGCTCTGTAAGGAGCA 1440  
DB 1381 AGGGTCTTTTCCCTCTCGCCAAAGAAATGCAAGTCTGTTGAATGCTCTGTAAGGAGCA 1440  
QY 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACAAAGTCTGTGAGCACCTTTGACGAGCGG 1500  
DB 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACAAAGTCTGTGAGCACCTTTGACGAGCGG 1500

QY 1501 AACCCCCACCTGGCGACAGGTGCTCTCGGCCCCAAAGCCACGTTGTAAGATACACT 1560  
DB 1501 AACCCCCACCTGGCGACAGGTGCTCTCGGCCCCAAAGCCACGTTGTAAGATACACT 1560  
QY 1561 GCAAGGCGGACAAACCCAGTGCACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAAA 1620  
DB 1561 GCAAGGCGGACAAACCCAGTGCACGTTGTGAGTTGGATAGTTGTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAGAGGTACCCCATGT 1680  
DB 1621 TGGCTCTCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAGAGGTACCCCATGT 1680  
QY 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTTACATGTTTGTGCGAGTTAAAA 1740  
DB 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTTACATGTTTGTGCGAGTTAAAA 1740  
QY 1741 AACGCTTAGGCGCCCCCGAAACACGGGAGCGTGTTCCTTTTGAAAAACAGATAATACC 1800  
DB 1741 AACGCTTAGGCGCCCCCGAAACACGGGAGCGTGTTCCTTTTGAAAAACAGATAATACC 1800  
QY 1801 ATGGCGCTATTACGGCTACTCCAAACAGACGGGAGGCTTCAAGTGTCTCCACCGCA 1860  
DB 1801 ATGGCGCTATTACGGCTACTCCAAACAGACGGGAGGCTTCAAGTGTCTCCACCGCA 1860  
QY 1861 AGCTCTACAGGCGGGACAGGAACAGGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
DB 1861 AGCTCTACAGGCGGGACAGGAACAGGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
QY 1921 ACACATCTTTCTGGGACCTGGTCAATGGGCTGTGTTGGACTGTCTATCATGTGGCC 1980  
DB 1921 ACACATCTTTCTGGGACCTGGTCAATGGGCTGTGTTGGACTGTCTATCATGTGGCC 1980  
QY 1981 GGCTCAAGACCTTTGCGGCCCCAAAGGGGCCCAATCAACCAAAATGTACCAATGTGGAC 2040  
DB 1981 GGCTCAAGACCTTTGCGGCCCCAAAGGGGCCCAATCAACCAAAATGTACCAATGTGGAC 2040  
QY 2041 CAGGACTCTGTGGCTGGCAAGCGCCCCCGGGGCGGTTCTTGACACCATGACCTGCG 2100  
DB 2041 CAGGACTCTGTGGCTGGCAAGCGCCCCCGGGGCGGTTCTTGACACCATGACCTGCG 2100  
QY 2101 GGCGACTCGGACCTTTACTTGGTTCAGAGGATGCCGATGTCTTCGGTGCCTGGCGG 2160  
DB 2101 GGCGACTCGGACCTTTACTTGGTTCAGAGGATGCCGATGTCTTCGGTGCCTGGCGG 2160  
QY 2161 GCGACAGCAGGGGAGCTTCTCTCCCGAGCGCTCTCTACTTTGAAGGCTCTTTCG 2220  
DB 2161 GCGACAGCAGGGGAGCTTCTCTCCCGAGCGCTCTCTCTCCCGAGCGCTCTCTACTTTGAAGGCTCTTTCG 2220  
QY 2221 GGGGCTCCTGCTCTGCTCGGCGACGCTGTGGGCACTTTTCGGGCTGCGGCTGTGC 2280  
DB 2221 GGGGCTCCTGCTCTGCTCGGCGACGCTGTGGGCACTTTTCGGGCTGCGGCTGTGC 2280  
QY 2281 ACCCGAGGGGTTGCGAAGCGGTGACTTTGTACCGTCTGAGTCTATGAAACCACTATG 2340  
DB 2281 ACCCGAGGGGTTGCGAAGCGGTGACTTTGTACCGTCTGAGTCTATGAAACCACTATG 2340  
QY 2341 CGGTCCCGGCTTTCACGGACAACTCGTCCCTCGCGCGTACCGCAGACATTCAGGTTG 2400  
DB 2341 CGGTCCCGGCTTTCACGGACAACTCGTCCCTCGCGCGTACCGCAGACATTCAGGTTG 2400  
QY 2401 GCCCATCTACGCGGCTTACTGTTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
DB 2401 GCCCATCTACGCGGCTTACTGTTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
QY 2461 GCGGAGGATTAAGGTGCTTGTCTGAACCGCTCGGCGCGCACCTAGTTTCGGG 2520  
DB 2461 GCGGAGGATTAAGGTGCTTGTCTGAACCGCTCGGCGCGCACCTAGTTTCGGG 2520  
QY 2521 GGTATATGTCTAAGGACATGTTATCGACCTTAACATCAGAACCGGGTAAGGACCATC 2580  
DB 2521 GGTATATGTCTAAGGACATGTTATCGACCTTAACATCAGAACCGGGTAAGGACCATC 2580  
QY 2581 ACCAGGGTGGCCCATCATCAGCTACTCCACCTATGCAAGTTTCTTGGCGGCGGTGGTTCG 2640

[illegible]



QY 4801 CAAACCACTGCCCCATGTGGAGCACAGATCAACGACATGTGAAAAACGGTTCCATGAGG 4860  
DB 4801 CAAACCACTGCCCCATGTGGAGCACAGATCAACGACATGTGAAAAACGGTTCCATGAGG 4860  
QY 4861 ATCGTGGGGCCCTAGAACCTGTAGTAACAGTGGGATGGAACATTCGCCATTAAACCGGTAC 4920  
DB 4861 ATCGTGGGGCCCTAGAACCTGTAGTAACAGTGGGATGGAACATTCGCCATTAAACCGGTAC 4920  
QY 4921 ACCACGGGCCCTGACAGCCCTCCCGGGGCGCCAAATTATTCTAGGCGGTGTGGGGGTG 4980  
DB 4921 ACCACGGGCCCTGACAGCCCTCCCGGGGCGCCAAATTATTCTAGGCGGTGTGGGGGTG 4980  
QY 4981 GCTGCTGAGGAGTACGTGGAGGTATACGGGGGTGGGGGATTTCCACTAGCTAGCGGGCATG 5040  
DB 4981 GCTGCTGAGGAGTACGTGGAGGTATACGGGGGTGGGGGATTTCCACTAGCTAGCGGGCATG 5040  
QY 5041 ACCACTGACACGTMAAGTGCCCGTGTAGTTCGGGCCCGCGAATTTCTTACAGAGTG 5100  
DB 5041 ACCACTGACACGTMAAGTGCCCGTGTAGTTCGGGCCCGCGAATTTCTTACAGAGTG 5100  
QY 5101 GATGGGTGGGGTGGACAGAGTACGCTCCAGGGTGCAACCCCTCTTACGGGAGAGTGC 5160  
DB 5101 GATGGGTGGGGTGGACAGAGTACGCTCCAGGGTGCAACCCCTCTTACGGGAGAGTGC 5160  
QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGCTGGGTACAGCTCCCATGCGAGCCGAA 5220  
DB 5161 ACATTCCTGCTGGGCTCAATCAATACCTGCTGGGTACAGCTCCCATGCGAGCCGAA 5220  
QY 5221 CCGGACGTAGAGTGTCACTTCCATGCTACCGACCCCTCCACATTAACGGCGGAGACG 5280  
DB 5221 CCGGACGTAGAGTGTCACTTCCATGCTACCGACCCCTCCACATTAACGGCGGAGACG 5280  
QY 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTGGCCAGCTCATAGCTAGCGAG 5340  
DB 5281 GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTGGCCAGCTCATAGCTAGCGAG 5340  
QY 5341 CTGCTGCGGCTTCCTTGAAGGCAACATGCACTACCGTCACTGCTACGCTCCCGGAGCTGAC 5400  
DB 5341 CTGCTGCGGCTTCCTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCTGAC 5400  
QY 5401 CTGCTGCGGCTTCCTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCTGAC 5460  
DB 5401 CTGCTGCGGCTTCCTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCTGAC 5460  
QY 5461 TCAGAAAAAAGTGTGTAATTTGGACTCTTTCGAGCCGTCCAAAGCGGAGGATGAG 5520  
DB 5461 TCAGAAAAAAGTGTGTAATTTGGACTCTTTCGAGCCGTCCAAAGCGGAGGATGAG 5520  
QY 5521 AGGGAAGTATCCGTTCCGGCGGAGATCCTCGGCGGAGTCCAGGAATTCCTCGAGCGATG 5580  
DB 5521 AGGGAAGTATCCGTTCCGGCGGAGATCCTCGGCGGAGTCCAGGAATTCCTCGAGCGATG 5580  
QY 5581 CCATATGGGCGCGCCGGATTAACACCTTCACTGTAGAGTCTCTGGAAGGACCCCGAC 5640  
DB 5581 CCATATGGGCGCGCCGGATTAACACCTTCACTGTAGAGTCTCTGGAAGGACCCCGAC 5640  
QY 5641 TAGTCCCTCCAGTGTGTAACAGGCTGTGATTCGCGCTGCGAAGCCCTCCGATPACCA 5700  
DB 5641 TAGTCCCTCCAGTGTGTAACAGGCTGTGATTCGCGCTGCGAAGCCCTCCGATPACCA 5700  
QY 5701 CCTTCCAGGAGGAGGAGCGGTGTGCTGTGAGATCTTACCGTGTCTTCTGCTTGGG 5760  
DB 5701 CCTTCCAGGAGGAGGAGCGGTGTGCTGTGAGATCTTACCGTGTCTTCTGCTTGGG 5760  
QY 5761 GAGCTGCCACAAAGACCTTCGGACGCTCCGATCGTCCGCGGTGACAGCGGCA 5820  
DB 5761 GAGCTGCCACAAAGACCTTCGGACGCTCCGATCGTCCGCGGTGACAGCGGCA 5820  
QY 5821 ACGGCTCTCTTACAGCCCTCCGAGCGAGCGGATCCGAGCTGAGTGTGAC 5880  
DB 5821 ACGGCTCTCTTACAGCCCTCCGAGCGAGCGGATCCGAGCTGAGTGTGAC 5880

QY 5881 TCCTCCATGCCCCCTCTTGAAGGAGGAGCGGGGATCCCGATCTCAGCGACGGTCTTGG 5940  
DB 5881 TCCTCCATGCCCCCTCTTGAAGGAGGAGCGGGGATCCCGATCTCAGCGACGGTCTTGG 5940  
QY 5941 TCTACCGTAAGCGGAGGAGCTAGTAGGACGCTCGTCTGCTGCTGATGCTTACACATGG 6000  
DB 5941 TCTACCGTAAGCGGAGGAGCTAGTAGGACGCTCGTCTGCTGCTGATGCTTACACATGG 6000  
QY 6001 ACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAACCAAGCTGCCCATCAATCACTG 6060  
DB 6001 ACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAACCAAGCTGCCCATCAATCACTG 6060  
QY 6061 AGCAACTCTTTGCTCCGTCACCAACTTGTCTATGCTACCAACTCTCGCAGCGCAAGC 6120  
DB 6061 AGCAACTCTTTGCTCCGTCACCAACTTGTCTATGCTACCAACTCTCGCAGCGCAAGC 6120  
QY 6121 CTGCGGAGAGAAAGGTCACCTTTGACAGACTGAGGCTCTGGAGGACCACTACCGGAC 6180  
DB 6121 CTGCGGAGAGAAAGGTCACCTTTGACAGACTGAGGCTCTGGAGGACCACTACCGGAC 6180  
QY 6181 GTGCTCAAGGAGATGAAAGGCGAAGCGCTCAAGTTAAGGCTAAACTTCTATCCGTGAG 6240  
DB 6181 GTGCTCAAGGAGATGAAAGGCGAAGCGCTCAAGTTAAGGCTAAACTTCTATCCGTGAG 6240  
QY 6241 GAAGCTGTAAAGTGAAGCGCCCACTTCGGCCAGATCTAAATTTGGCTATGGGCGAAAG 6300  
DB 6241 GAAGCTGTAAAGTGAAGCGCCCACTTCGGCCAGATCTAAATTTGGCTATGGGCGAAAG 6300  
QY 6301 GACGTCCGGAACCTATCCAGCAAGCGGTTAAACCACTCCGCTCGGTGTGGAAGACTG 6360  
DB 6301 GACGTCCGGAACCTATCCAGCAAGCGGTTAAACCACTCCGCTCGGTGTGGAAGACTG 6360  
QY 6361 CTGGAAGACTGAGACACCAATTTGACACCACTCATGGAAGGAGGTTTCTGCG 6420  
DB 6361 CTGGAAGACTGAGACACCAATTTGACACCACTCATGGAAGGAGGTTTCTGCG 6420  
QY 6421 GTCCAAACGAGAGGAGGCGCGCAAGCCAGCTCGCTTATCGTATTTCCAGATTTGGG 6480  
DB 6421 GTCCAAACGAGAGGAGGCGCGCAAGCCAGCTCGCTTATCGTATTTCCAGATTTGGG 6480  
QY 6481 GTTCGTGTGTCGAGAAATGGCCCTTACGATGTGTCTCCACCTTCCAGCGCGTG 6540  
DB 6481 GTTCGTGTGTCGAGAAATGGCCCTTACGATGTGTCTCCACCTTCCAGCGCGTG 6540  
QY 6541 ATGGGCTCTTCACTACCGATTTCCAAATCTCTCTGGACAGCGGCTCGAGTTCGTGAAT 6600  
DB 6541 ATGGGCTCTTCACTACCGATTTCCAAATCTCTCTGGACAGCGGCTCGAGTTCGTGAAT 6600  
QY 6601 GCTGGAAGCGGAGAAATGCCCTATGGGCTTCGCATATGACACCCGCTGTTTGACTCA 6660  
DB 6601 GCTGGAAGCGGAGAAATGCCCTATGGGCTTCGCATATGACACCCGCTGTTTGACTCA 6660  
QY 6661 ACGETCACTGAGATGACATCGTGTGAGGAGTCAATCTTACCAATGTTGTGACTTGGCC 6720  
DB 6661 ACGETCACTGAGATGACATCGTGTGAGGAGTCAATCTTACCAATGTTGTGACTTGGCC 6720  
QY 6721 CCGAAGCCAGACAGGCCATAAGGTCGCTCACAGAGCGGCTTACATCGGGGCGCCCTG 6780  
DB 6721 CCGAAGCCAGACAGGCCATAAGGTCGCTCACAGAGCGGCTTACATCGGGGCGCCCTG 6780  
QY 6781 ACTAATTTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGCGGTGTACTGAGC 6840  
DB 6781 ACTAATTTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGCGGTGTACTGAGC 6840  
QY 6841 ACCAGTCCGAGTAAATACCTTCATGTTTAAAGGCGGCTGCGGCTGTGAGCTGCG 6900  
DB 6841 ACCAGTCCGAGTAAATACCTTCATGTTTAAAGGCGGCTGCGGCTGTGAGCTGCG 6900  
QY 6901 AAGCTCCAGGACTGACGATGCTTGTGCGGAGAGGAGCTTGTGTTATCTGTGAAGC 6960  
DB 6901 AAGCTCCAGGACTGACGATGCTTGTGCGGAGAGGAGCTTGTGTTATCTGTGAAGC 6960  
QY 6961 GCGGGAGCCCAAGAGGAGGAGCGGAGCGCTACGGGCGCTTTCAGCGAGGCTATGACTAGATAC 7020



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 20:21:40 ; Search time 564 Seconds  
(without alignments)  
10068.249 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989  
Sequence: 1 gccagcccccagttgggggc.....ggcctctctgcagatacaagt 7989

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 35539441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	7987.4	100.0	7989	4	US-09-539-601-10
2	7965.4	99.7	8001	4	US-09-539-601-7
3	7947.8	99.5	8001	4	US-09-539-601-22
4	7946.2	99.5	8001	4	US-09-539-601-16
5	7939.8	99.4	8001	4	US-09-539-601-28
6	7329.4	91.7	8637	4	US-09-539-601-4
7	7307.4	91.5	8649	4	US-09-539-601-13
8	7272.2	91.0	8639	4	US-10-029-907-1
9	7271.2	91.0	8638	4	US-10-029-907-24
10	7248.2	90.7	8638	4	US-10-029-907-7
11	7247.2	90.7	8638	4	US-10-029-907-25
12	7245.2	90.7	8642	4	US-10-029-907-2
13	7244	90.7	8638	4	US-10-029-907-6
14	7228.2	90.5	8648	4	US-10-029-907-5
15	7221	90.4	8643	4	US-10-029-907-4
16	6185.2	77.4	11076	4	US-09-539-601-1
17	6167.6	77.2	11076	4	US-09-539-601-25
18	6166	77.2	11076	4	US-09-539-601-19
19	6159.6	77.1	11076	4	US-09-539-601-31
20	5316.8	66.6	9595	3	US-09-014-416-4
21	5216.4	65.3	7917	1	US-08-324-977-31
22	5216.4	65.3	7917	2	US-08-384-616-31
23	5216.4	65.3	7917	2	US-08-904-686A-31
24	5216.4	65.3	7917	3	US-09-315-850-31
25	5216.4	65.3	9416	1	US-08-324-977-1
26	5216.4	65.3	9416	2	US-08-384-616-1
27	5216.4	65.3	9416	2	US-08-904-686A-1

28 5216.4 65.3 9416 3 US-09-315-850-1  
29 5216.4 65.3 9416 4 US-08-823-895A-27  
30 5210.2 65.2 9472 4 US-08-150-204B-96  
31 5175.2 64.8 7863 1 US-08-324-977-35  
32 5175.2 64.8 7863 2 US-08-384-616-35  
33 5175.2 64.8 7863 2 US-08-904-686A-35  
34 5175.2 64.8 7863 3 US-09-315-850-35  
35 5175.2 64.8 9030 1 US-08-324-977-13  
36 5175.2 64.8 9030 2 US-08-384-616-13  
37 5175.2 64.8 9030 3 US-08-904-686A-13  
38 5175.2 64.8 9030 2 US-09-315-850-13  
39 4102.2 51.3 9599 3 US-09-014-416-2  
40 4100.6 51.3 9599 3 US-09-014-416-6  
41 4096.2 51.3 9646 3 US-08-811-566-1  
42 4096.2 51.3 9646 3 US-09-034-756-1  
43 4092.6 51.2 12980 3 US-08-811-566-5  
44 4092.6 51.2 12980 3 US-09-034-756-5  
45 4019.8 50.3 9379 3 US-09-388-874-1

#### ALIGNMENTS

RESULT 1  
US-09-539-601-10  
; Sequence 10, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 10  
; LENGTH: 7989  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I377/NS3-3'/wt  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1181)  
; OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase  
; OTHER INFORMATION: fusion protein  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (1190)..(1800)  
; OTHER INFORMATION: internal ribosome entry site from  
; OTHER INFORMATION: encephalomyocarditis virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1801)..(7758)  
; OTHER INFORMATION: hepatitis C virus NS3 - 5B  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (7759)..(7989)  
; PUBLICATION INFORMATION:  
; AUTHORS: Lohmann, Volker  
; AUTHORS: Krner, Frank  
; AUTHORS: Koch, Jan-Oliver  
; AUTHORS: Herian, Ulrike  
; AUTHORS: Theilmann, Lorenz  
; AUTHORS: Bartenschlager, Ralf  
; TITLE: Replication of subgenomic hepatitis C virus RNAs in a  
; TITLE: Hepatoma cell line  
; JOURNAL: Science  
; VOLUME: 285

Sequence 1, Appli  
Sequence 27, Appl  
Sequence 96, Appl  
Sequence 35, Appl  
Sequence 35, Appl  
Sequence 35, Appl  
Sequence 35, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 6, Appl  
Sequence 6, Appl  
Sequence 1, Appl  
Sequence 1, Appl  
Sequence 5, Appl  
Sequence 5, Appl  
Sequence 1, Appl

;  
; PAGES: 110-113  
; DATE: 1999-07-02  
; US-09-539-601-10

Query Match									
Best Local Similarity 100.0%; Score 7987.4; DB 4; Length 7989;									
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;									
QY	1	GCACGCCCCGATGGGGGACACTCCACATAGATCACTCCCTCTGAGGAACACTAG	60						
DB	1	GCACGCCCCGATGGGGGACACTCCACATAGATCACTCCCTCTGAGGAACACTAG	60						
QY	61	TCCTTCAGCAGAAAGCGCTTAGCCATCGCGCTAGTATGAGTGTGCTCAGCCTCCAGAC	120						
DB	61	TCCTTCAGCAGAAAGCGCTTAGCCATCGCGCTAGTATGAGTGTGCTCAGCCTCCAGAC	120						
QY	121	CCCCCTCCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTCGAC	180						
DB	121	CCCCCTCCCGGAGAGCCATAGTGTCTCGGAACCGGTGAGTACACCGGAATTCGAC	180						
QY	181	GACGACCGGTCTCTTCTTGATCAACCCCGCTCAATGCTGGAGATTGGCGGTGCCCC	240						
DB	181	GACGACCGGTCTCTTCTTGATCAACCCCGCTCAATGCTGGAGATTGGCGGTGCCCC	240						
QY	241	GCAGAGCTGCTAGCCGAGTAGTGTGGTTCGGAAGCGCTTGTGTAATTCGCTGATAG	300						
DB	241	GCAGAGCTGCTAGCCGAGTAGTGTGGTTCGGAAGCGCTTGTGTAATTCGCTGATAG	300						
QY	301	GTGCTTGCAGTGCCTCGGAGTCTGTGAGACCGTGCACATGAGCACGAATCCTAAAC	360						
DB	301	GTGCTTGCAGTGCCTCGGAGTCTGTGAGACCGTGCACATGAGCACGAATCCTAAAC	360						
QY	361	CTCAAGAAACCAAGGCGCGCATGATTCAAGATGATTCGACGAGGTCTC	420						
DB	361	CTCAAGAAACCAAGGCGCGCATGATTCAAGATGATTCGACGAGGTCTC	420						
QY	421	CGCCCGCTTGGTGGAGAGCTATTGGCTATGACTGGGCACACAGACATCGGTGCT	480						
DB	421	CGCCCGCTTGGTGGAGAGCTATTGGCTATGACTGGGCACACAGACATCGGTGCT	480						
QY	481	CTGATGCGCGGTTCCTCGGCTGTGAGCAGGGCGCCGGTCTTTTGTCAAGACCG	540						
DB	481	CTGATGCGCGGTTCCTCGGCTGTGAGCAGGGCGCCGGTCTTTTGTCAAGACCG	540						
QY	541	ACCTGTCCGTGCTGATGAATGAATGAACTGAGACGAGGCGCGGTATCGTCTGCCA	600						
DB	541	ACCTGTCCGTGCTGATGAATGAACTGAGACGAGGCGCGGTATCGTCTGCCA	600						
QY	601	CGACGGCGTTCCTTGCAGCTGTGCTGACGCTTGTCACTGAAGCGGGAAGGACTGGC	660						
DB	601	CGACGGCGTTCCTTGCAGCTGTGCTGACGCTTGTCACTGAAGCGGGAAGGACTGGC	660						
QY	661	TGCTATTGGCGAGTGCCTGGGCGAGATCTCTGTCACTCTACCTTGTCTCTGCCGAG	720						
DB	661	TGCTATTGGCGAGTGCCTGGGCGAGATCTCTGTCACTCTACCTTGTCTCTGCCGAG	720						
QY	721	AAGTATCCATCATGGCTGATGCAATCGCGGCTGCATACGCTTGTATCCGGTACCTGCC	780						
DB	721	AAGTATCCATCATGGCTGATGCAATCGCGGCTGCATACGCTTGTATCCGGTACCTGCC	780						
QY	781	CATTGACCAACCAAGCGAACAATCGCATCGAGCGACAGTATCTCGGATGGAAGCGGTC	840						
DB	781	CATTGACCAACCAAGCGAACAATCGCATCGAGCGACAGTATCTCGGATGGAAGCGGTC	840						
QY	841	TTGTGATCAGATGATCTGGAGAGAGATCAGGGGTCTGGCGAGCGGAATCTGTTTCG	900						
DB	841	TTGTGATCAGATGATCTGGAGAGAGATCAGGGGTCTGGCGAGCGGAATCTGTTTCG	900						
QY	901	CCAGGCTCAAGGCGGATGCTCCGAGCGGAGGATCTCGTGTGACCCATGCGATGCTCT	960						
DB	901	CCAGGCTCAAGGCGGATGCTCCGAGCGGAGGATCTCGTGTGACCCATGCGATGCTCT	960						
QY	961	GCTTCCCGAATATCATGGTGGAAATGGCGGCTTTTCTGGATTCATCGACTGTGGCGGC	1020						

DB	961	GCTTCCCGAATATCATGGTGGAAATGGCGGCTTTTCTGGATTCATCGACTGTGGCGGC	1020						
QY	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATTGCTGAAGAGC	1080						
DB	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATTGCTGAAGAGC	1080						
QY	1081	TTGGCGGGAATGGGTGACCGCTTCTCTGTCTTACGGTATCGCGCTCCCGATTGCG	1140						
DB	1081	TTGGCGGGAATGGGTGACCGCTTCTCTGTCTTACGGTATCGCGCTCCCGATTGCG	1140						
QY	1141	AGCGCATCGCTTCTATCGCTTCTTGAACGAGTCTTCTGAGTTTAAACAGACACCAACG	1200						
DB	1141	AGCGCATCGCTTCTATCGCTTCTTGAACGAGTCTTCTGAGTTTAAACAGACACCAACG	1200						
QY	1201	GTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTAACTACTTGGC	1260						
DB	1201	GTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTAACTACTTGGC	1260						
QY	1261	CGAAGCGCTTGAATAAGCGCGGTGTGGTGTCTATATGTTATTTCCACCATATTG	1320						
DB	1261	CGAAGCGCTTGAATAAGCGCGGTGTGGTGTCTATATGTTATTTCCACCATATTG	1320						
QY	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTCTTGAAGCATTCCT	1380						
DB	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGGCCCTGTCTTCTTGAAGCATTCCT	1380						
QY	1381	AGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTGTGAAGGAACA	1440						
DB	1381	AGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTGTGAAGGAACA	1440						
QY	1441	GTTCCTCTGGAGCTTCTTGAAGACAAACAACTGTGTAGCGACCTTTTSCAGGACGG	1500						
DB	1441	GTTCCTCTGGAGCTTCTTGAAGACAAACAACTGTGTAGCGACCTTTTSCAGGACGG	1500						
QY	1501	AACCCCGACCTTGGGACAGGTCCCTCTCGGCCAAAGACAGCTGTATAAGATACACCT	1560						
DB	1501	AACCCCGACCTTGGGACAGGTCCCTCTCGGCCAAAGACAGCTGTATAAGATACACCT	1560						
QY	1561	GCAAGCGGCAACAAACCCAGTGCACCTTGTGAGTTGGATAGTTGGAAAGAGTCAAA	1620						
DB	1561	GCAAGCGGCAACAAACCCAGTGCACCTTGTGAGTTGGATAGTTGGAAAGAGTCAAA	1620						
QY	1621	TGGCTCTCTCAAGCTATTCAAGGGGCTGAAGGATGCCAGAAAGTACCCATTGT	1680						
DB	1621	TGGCTCTCTCAAGCTATTCAAGGGGCTGAAGGATGCCAGAAAGTACCCATTGT	1680						
QY	1681	ATGGGATCTGATCTGGGCTCGGTGCACATGCTTTTACATGCTTTAGTCGAGTTTAAA	1740						
DB	1681	ATGGGATCTGATCTGGGCTCGGTGCACATGCTTTTACATGCTTTAGTCGAGTTTAAA	1740						
QY	1741	AACGTCTAGGCCCCCGAACACCGGACCGTGGTTTCTTTTGAATAACACGATAATAC	1800						
DB	1741	AACGTCTAGGCCCCCGAACACCGGACCGTGGTTTCTTTTGAATAACACGATAATAC	1800						
QY	1801	ATGGCGCTTATTACGCTTCTCCCAAGAGCGGAGCTTACTTGGCTGCATCATCT	1860						
DB	1801	ATGGCGCTTATTACGCTTCTCCCAAGAGCGGAGCTTACTTGGCTGCATCATCT	1860						
QY	1861	AGCCTCACAGGCGGACAGGAACAGGTGAGGGGAGGTCCAGTGTCTCCACCGCA	1920						
DB	1861	AGCCTCACAGGCGGACAGGAACAGGTGAGGGGAGGTCCAGTGTCTCCACCGCA	1920						
QY	1921	ACACATCTTCTCGCGACCTTGGTCAATGGCGTGTGTGACTGTCTATCATGTGTC	1980						
DB	1921	ACACATCTTCTCGCGACCTTGGTCAATGGCGTGTGTGACTGTCTATCATGTGTC	1980						
QY	1981	GGCTCAAGACCTTTCGCGGCCCCAAAGGGCCCCAACATGACCAATGTGACCAATGTGGAC	2040						
DB	1981	GGCTCAAGACCTTTCGCGGCCCCAAAGGGCCCCAACATGACCAATGTGACCAATGTGGAC	2040						
QY	2041	CAGGACCTCGTGGTGGCAAGCGCCCCGGGCGGTTCCTTGACCATGACCTGCTGC	2100						

Db 2041 CAGGACCTCGTCGGCTGGCAAGCGCCGCCCGCGGTCCTCTTGACACCATGACACCTGC 2100  
Qy  
2101 GCGAGCTCGGACCTTACTTGGTACAGAGGATGCGGATGTCATTCGGTCCGCGCGG 2160  
Db  
2101 GCGAGCTCGGACCTTACTTGGTACAGAGGATGCGGATGTCATTCGGTCCGCGCGG 2160  
Qy  
2161 GCGGACAGCAGGGGAGCCTACTCTCCGCCAGGCGGTCCTCTACTTGAAGGGCTCTTCG 2220  
Db  
2161 GCGGACAGCAGGGGAGCCTACTCTCCGCCAGGCGGTCCTCTACTTGAAGGGCTCTTCG 2220  
Qy  
2221 GCGGTCACCTGCTCTGCCCTCGGGGACAGCTGTGGGCATCTTTTGGGGCTGCGGTGTGC 2280  
Db  
2221 GCGGTCACCTGCTCTGCCCTCGGGGACAGCTGTGGGCATCTTTTGGGGCTGCGGTGTGC 2280  
Qy  
2281 ACCCGAGGGGTGGAGGGCGGTGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
Db  
2281 ACCCGAGGGGTGGAGGGCGGTGACTTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
Qy  
2341 CGGTCCCGGTCTTTCAGGCAAACTCGTCCCTCCGGCGGTACCGCAGACATTCAGGTG 2400  
Db  
2341 CGGTCCCGGTCTTTCAGGCAAACTCGTCCCTCCGGCGGTACCGCAGACATTCAGGTG 2400  
Qy  
2401 GCCCATCTACAGCCCTACTGTGAGGGCAAGAGCACTAAGTGCCTGGCTGCGTATGCA 2460  
Db  
2401 GCCCATCTACAGCCCTACTGTGAGGGCAAGAGCACTAAGTGCCTGGCTGCGTATGCA 2460  
Qy  
2461 GCCCAAGGGTATAGGTGCTGTCTGACCCGTCCGTGCGCGGACCGCTAGGTTCGGG 2520  
Db  
2461 GCCCAAGGGTATAGGTGCTGTCTGACCCGTCCGTGCGCGGACCGCTAGGTTCGGG 2520  
Qy  
2521 GCGTATATGCTTAAGGCACATGTTATCGACCTAACATCAGAACCGGGTAAGGACCATC 2580  
Db  
2521 GCGTATATGCTTAAGGCACATGTTATCGACCTAACATCAGAACCGGGTAAGGACCATC 2580  
Qy  
2581 ACCACGGGTGCCCCATCACTACTCCACCTATGSCAAAGTTCTTGGCGAGCGGTGTC 2640  
Db  
2581 ACCACGGGTGCCCCATCACTACTCCACCTATGSCAAAGTTCTTGGCGAGCGGTGTC 2640  
Qy  
2641 TCTGGGGCGCCTATGACATCATATATGATGAGTGCGACTCAACTGACCTGACCACT 2700  
Db  
2641 TCTGGGGCGCCTATGACATCATATATGATGAGTGCGACTCAACTGACCTGACCACT 2700  
Qy  
2701 ATCTGGGCTCGGCACAGTCTCGACCAAGCGGAGCGGTGAGCGGACTGCTGCTG 2760  
Db  
2701 ATCTGGGCTCGGCACAGTCTCGACCAAGCGGAGCGGTGAGCGGACTGCTGCTG 2760  
Qy  
2761 CTGCCACCGCTAGCGCTCGGGATCGGTCAACCGTGCCATCCAAACATCGAGAGGTG 2820  
Db  
2761 CTGCCACCGCTAGCGCTCGGGATCGGTCAACCGTGCCATCCAAACATCGAGAGGTG 2820  
Qy  
2821 GCTCTGTCCAGCTGAGAAATCCCTTTATGGGAAAGCAATCCCAATCGAGACCATC 2880  
Db  
2821 GCTCTGTCCAGCTGAGAAATCCCTTTATGGGAAAGCAATCCCAATCGAGACCATC 2880  
Qy  
2881 AAGGGGGAGGACCTCACTTTCTGCCATTCGAAGAAAGATGATGAGTCCGCGG 2940  
Db  
2881 AAGGGGGAGGACCTCACTTTCTGCCATTCGAAGAAAGATGATGAGTCCGCGG 2940  
Qy  
2941 AAGCTGTCCGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTGTATGATCGTC 3000  
Db  
2941 AAGCTGTCCGCTCGGACTCAATGTGTAGCATATTTACCGGGGCTTGTATGATCGTC 3000  
Qy  
3001 ATACCACTAGCGAGAGCTCATGTGTAGCAACGAGCGCTCATAGCGGCTTTACC 3060  
Db  
3001 ATACCACTAGCGAGAGCTCATGTGTAGCAACGAGCGCTCATAGCGGCTTTACC 3060  
Qy  
3061 GCGGATTTGACTCAGTGATCGACTCAATGATGTCACCCAGACAGTCGACTTCAGC 3120  
Db  
3061 GCGGATTTGACTCAGTGATCGACTCAATGATGTCACCCAGACAGTCGACTTCAGC 3120  
Qy  
3121 CTGACCCGACCTTACCAATTGAGACGACGACCGGTGCGCAAGACGCGGTGTACGCTG 3180  
Db  
3121 CTGACCCGACCTTACCAATTGAGACGACGACCGGTGCGCAAGACGCGGTGTACGCTG 3180

Qy 3181 CAGCGGCGAGGACGAGCTGTGTAGGGGAGAGATGGGCACTTTACAGTTTGTGACTCCAGA 3240  
Db  
3181 CAGCGGCGAGGACGAGCTGTGTAGGGGAGAGATGGGCACTTTACAGTTTGTGACTCCAGA 3240  
Qy  
3241 GAAAGCGCCCTCGGACATGTCGATTCTCGTTCTGTCGAGTGTCTATGACGCGGCTGT 3300  
Db  
3241 GAAAGCGCCCTCGGACATGTCGATTCTCGTTCTGTCGAGTGTCTATGACGCGGCTGT 3300  
Qy  
3301 GCTTGTGTACGAGCTCAGCGCCCGGAGACCTCAGTTAGTTTGGGGCTTACTCTAAACACA 3360  
Db  
3301 GCTTGTGTACGAGCTCAGCGCCCGGAGACCTCAGTTAGTTTGGGGCTTACTCTAAACACA 3360  
Qy  
3361 CGAGGTTGCCCTGTCGAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCGTC 3420  
Db  
3361 CGAGGTTGCCCTGTCGAGGACCATCTGGAGTTCTGGGAGAGCGTCTTTACAGGCGTC 3420  
Qy  
3421 ACCCACATAGACGCCCATTTCTGTCCAGACTAAGCAGGAGAGACACTTCCCTTAC 3480  
Db  
3421 ACCCACATAGACGCCCATTTCTGTCCAGACTAAGCAGGAGAGAGACACTTCCCTTAC 3480  
Qy  
3481 CTGTGTACATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCTCCATCTGTGGAC 3540  
Db  
3481 CTGTGTACATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCTCCATCTGTGGAC 3540  
Qy  
3541 CAATGTGGAGTGTCTCATACGCTAAAGCTTAAGCTGCAAGGGGCAACGCCCTGCTG 3600  
Db  
3541 CAATGTGGAGTGTCTCATACGCTAAAGCTTAAGCTGCAAGGGGCAACGCCCTGCTG 3600  
Qy  
3601 TATAGGCTGGAGCGGTTCAAAAGAGGTTACTACACACACCCCATTAACCAATATATC 3660  
Db  
3601 TATAGGCTGGAGCGGTTCAAAAGAGGTTACTACACACACCCCATTAACCAATATATC 3660  
Qy  
3661 ATGSCATGATGCTGCGCTGACCTGGAGTCTGTCACAGACACCTGGGTGCTGTAGGCGGA 3720  
Db  
3661 ATGSCATGATGCTGCGCTGACCTGGAGTCTGTCACAGACACCTGGGTGCTGTAGGCGGA 3720  
Qy  
3721 GTCTTAGACCTCTGGCGCGGTATGCTGACAAAGGAGCGGTGCTCATTTGGGAGG 3780  
Db  
3721 GTCTTAGACCTCTGGCGCGGTATGCTGACAAAGGAGCGGTGCTCATTTGGGAGG 3780  
Qy  
3781 ATCATCTTGTCCGAAAGCGGCATCATTTCCGACAGGAAAGTCTTTTACGGGAGTTC 3840  
Db  
3781 ATCATCTTGTCCGAAAGCGGCATCATTTCCGACAGGAAAGTCTTTTACGGGAGTTC 3840  
Qy  
3841 GATGAGATGGAAGATGCGCTCACTCTCCCTTACATCGAAAGGAAATGACCTGCC 3900  
Db  
3841 GATGAGATGGAAGATGCGCTCACTCTCCCTTACATCGAAAGGAAATGACCTGCC 3900  
Qy  
3901 GAACAAATTCAAAAGAGGCAATCGGGTTCGTCAAAACAGCCCAAGAGCGAGGCT 3960  
Db  
3901 GAACAAATTCAAAAGAGGCAATCGGGTTCGTCAAAACAGCCCAAGAGCGAGGCT 3960  
Qy  
3961 GCTGCTCCGCTGCTGGAATCCAGTGGCGGACCTCGAGCCTTCTGGGAGAGCATATG 4020  
Db  
3961 GCTGCTCCGCTGCTGGAATCCAGTGGCGGACCTCGAGCCTTCTGGGAGAGCATATG 4020  
Qy  
4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Db  
4021 TGGAAATTCATCAGCGGATACAAATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Qy  
4081 GCGATGATCATCTGATGAGCATTCAGGCTCTATCAACAGCCGCTTACACCCCAACAT 4140  
Db  
4081 GCGATGATCATCTGATGAGCATTCAGGCTCTATCAACAGCCGCTTACACCCCAACAT 4140  
Qy  
4141 ACCCTCTGTTTAAACATCTGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGGCT 4200  
Db  
4141 ACCCTCTGTTTAAACATCTGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGGCT 4200  
Qy  
4201 GCTTCTGCTTTCGTAGCGCGGCTATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260  
Db  
4201 GCTTCTGCTTTCGTAGCGCGGCTATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG 4260

4261 AAGTGTCTTGCGATATTTGGCAGGTATGAGCAGGGTGGCAGGCGCGCTCGTGGCC 4320  
4261 AAGTGTCTTGCGATATTTGGCAGGTATGAGCAGGGTGGCAGGCGCGCTCGTGGCC 4320  
4321 TTTAAGGTCAATGAGGGCGAGATGCCCTCCACCGAGACCTGGTTAAACCTACTCCCTGCT 4380  
4321 TTTAAGGTCAATGAGGGCGAGATGCCCTCCACCGAGACCTGGTTAAACCTACTCCCTGCT 4380  
4381 ATCCCTCTCCCTGGCGCCCTAGTCTGGGGTCTGTCGGCAGCGATATCGTGGGAC 4440  
4381 ATCCCTCTCCCTGGCGCCCTAGTCTGGGGTCTGTCGGCAGCGATATCGTGGGAC 4440  
4441 GTGGGCCACAGGGAGGGGCTGTGAGTGAATGAACCGGCTGATAGCGTTGCTTCGCG 4500  
4441 GTGGGCCACAGGGAGGGGCTGTGAGTGAATGAACCGGCTGATAGCGTTGCTTCGCG 4500  
4501 GGTAAACACGTCTCCCCACGCACTATGTGCTGAGAGCGCTGACAGCGTGTCACT 4560  
4561 CAGATCTCTAGTCTTACCATCACTAGCTGCTGAAGAGGTTCAACAGTGGATCAAC 4620  
4561 CAGATCTCTAGTCTTACCATCACTAGCTGCTGAAGAGGTTCAACAGTGGATCAAC 4620  
4621 GAGGACTGTCTCAGGCACTGCTCGGCTCGTGGCTAAGAGATGTTTGGGATGGATATGC 4680  
4621 GAGGACTGTCTCAGGCACTGCTCGGCTCGTGGCTAAGAGATGTTTGGGATGGATATGC 4680  
4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCCCGCGATTCGCGGA 4740  
4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCCCGCGATTCGCGGA 4740  
4741 GTCCCTCTTCTCATGTCAAAGTGGGTACGAGGAGTCTGGCGGGCGACGCACTATG 4800  
4741 GTCCCTCTTCTCATGTCAAAGTGGGTACGAGGAGTCTGGCGGGCGACGCACTATG 4800  
4801 CAAACCACTGCGCATGTGGAGCACAGATCACCGGACATGTGAAAAACGGTTCCATGAG 4860  
4801 CAAACCACTGCGCATGTGGAGCACAGATCACCGGACATGTGAAAAACGGTTCCATGAG 4860  
4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTCGCCATTAACGGTAC 4920  
4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTCGCCATTAACGGTAC 4920  
4921 ACCAGGGCCCTGACCGCTCCCGGCGCCAAATTTCTAGGGCGCTGTGGCGGTG 4980  
4921 ACCAGGGCCCTGACCGCTCCCGGCGCCAAATTTCTAGGGCGCTGTGGCGGTG 4980  
4981 GCTGCTGAGAGTACGTGAGGTTACGCGGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
4981 GCTGCTGAGAGTACGTGAGGTTACGCGGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
5041 ACCACTGACAAAGTAAAGTCCCGTGTGAGTTCCGGGCCCCCGAAATTCCTCAAGAAGTG 5100  
5041 ACCACTGACAAAGTAAAGTCCCGTGTGAGTTCCGGGCCCCCGAAATTCCTCAAGAAGTG 5100  
5101 GATGGGGTGGGTGACAGGTACGTCCAGCGTCAAAACCCCTCTACGGGAGGTC 5160  
5101 GATGGGGTGGGTGACAGGTACGTCCAGCGTCAAAACCCCTCTACGGGAGGTC 5160  
5161 ACATTCTGTCGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCCCGAA 5220  
5161 ACATTCTGTCGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCCCGAA 5220  
5221 CCGGAGTAGAGTGTCACTTCCATGCTACCGACCCCTCCACATTAACGGCGAGACG 5280  
5221 CCGGAGTAGAGTGTCACTTCCATGCTACCGACCCCTCCACATTAACGGCGAGACG 5280  
5281 GCTAAGGTAGGTGCGCAGGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
5281 GCTAAGGTAGGTGCGCAGGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG 5340  
5341 CTGTCTGCGCCTTCTTGAAGGCAACATGCACTACCGGTCAATGACTCCCGGACGCTGAC 5400

5341 CTGTCTGCGCCTTCTTGAAGGCAACATGCACTACCGGTCAATGACTCCCGGACGCTGAC 5400  
5401 CTCATCAGAGCCAACTCTGTGGCGGAGGAGATGGGGGAAATCAACCCCGTGGAG 5460  
5401 CTCATCAGAGCCAACTCTGTGGCGGAGGAGATGGGGGAAATCAACCCCGTGGAG 5460  
5461 TCAGAAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAAGCGAGGAGGATGAG 5520  
5461 TCAGAAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAAGCGAGGAGGATGAG 5520  
5521 AGGGAAGTATCCGTTCCGGCGAGATCTGTCGGAGGTCCAGAAAATTCCTCGAGGATG 5580  
5521 AGGGAAGTATCCGTTCCGGCGAGATCTGTCGGAGGTCCAGAAAATTCCTCGAGGATG 5580  
5581 CCCATATGGGACGCGCGGATTAACACCTCCACTGTGTAGAGTCTTGGAGGACCCGGAC 5640  
5581 CCCATATGGGACGCGCGGATTAACACCTCCACTGTGTAGAGTCTTGGAGGACCCGGAC 5640  
5641 TACGTCCCTCCAGTGTGATACACGGGTGTCATTTGCCGCTGCCAAAGCCCTCCGATACCA 5700  
5641 TACGTCCCTCCAGTGTGATACACGGGTGTCATTTGCCGCTGCCAAAGCCCTCCGATACCA 5700  
5701 CCTTCAAGGAGGAGGACGGTTGTCTGTCCAGAAATCTACCGTGTCTTCTGCTTGGG 5760  
5701 CCTTCAAGGAGGAGGACGGTTGTCTGTCCAGAAATCTACCGTGTCTTCTGCTTGGG 5760  
5761 GAGCTCGCCACAAAGACCTTCGGCAGCTCCGATCGTGGCGGTCCAGACGCGCACGGCA 5820  
5761 GAGCTCGCCACAAAGACCTTCGGCAGCTCCGATCGTGGCGGTCCAGACGCGCACGGCA 5820  
5821 ACGGCTCTCTTACAGCAGCCCTCCAGACGCGGACGCGGGGATCCGACGTTGAGTCTGATC 5880  
5821 ACGGCTCTCTTACAGCAGCCCTCCAGACGCGGACGCGGGGATCCGACGTTGAGTCTGATC 5880  
5881 TCTTCCATGCCCCCTTTCGAGGGGAGCGGGGATCCCGATCTCAGCGACGCGTCTTGG 5940  
5881 TCTTCCATGCCCCCTTTCGAGGGGAGCGGGGATCCCGATCTCAGCGACGCGTCTTGG 5940  
5941 TCTACGTAAGCGAGGAGCTAGTGAAGACGCTGCTGCTGCTGCTGCTGCTGCTACATGG 6000  
5941 TCTACGTAAGCGAGGAGCTAGTGAAGACGCTGCTGCTGCTGCTGCTGCTGCTACATGG 6000  
6001 ACAGGCGCTGATCAGCCATGCTGCGGAGGAAACCAAGCTGCCATCAATGACATG 6060  
6001 ACAGGCGCTGATCAGCCATGCTGCGGAGGAAACCAAGCTGCCATCAATGACATG 6060  
6061 AGCAACTCTTGTCCGTCCACCAACTTTGGTCTATGCTTACAACTCTCGAGCGCAAGC 6120  
6061 AGCAACTCTTGTCCGTCCACCAACTTTGGTCTATGCTTACAACTCTCGAGCGCAAGC 6120  
6121 CTGCGGACAGAGGAGTCACTTTGACAGCTGCGAGTCTCGACGACCACTACCGGAC 6180  
6121 CTGCGGACAGAGGAGTCACTTTGACAGCTGCGAGTCTCGACGACCACTACCGGAC 6180  
6181 GTGCTCAAGAGATGAAGCGGAGGCTCCACAGTTAAGGCTAAAATTCTATCCGTGGAG 6240  
6181 GTGCTCAAGAGATGAAGCGGAGGCTCCACAGTTAAGGCTAAAATTCTATCCGTGGAG 6240  
6241 GAAGCCTGTAAAGTGAAGCGGCGGAGATTCGCGGAGATCTTAAATTTGGCTATGGGGCAAG 6300  
6241 GAAGCCTGTAAAGTGAAGCGGCGGAGATTCGCGGAGATCTTAAATTTGGCTATGGGGCAAG 6300  
6301 GAGCTCCGAACTATCCAGAGGCGGTTAACCATCCGCTCCGTGTGGAGGAGGACTTG 6360  
6301 GAGCTCCGAACTATCCAGAGGCGGTTAACCATCCGCTCCGTGTGGAGGAGGACTTG 6360  
6361 CTGGAAGACACTGAGACACCAATTGACACCACTCATGCAAAATAAGGTTTCTGCG 6420  
6421 GTCCAAACGAGAGGGGGCGGCAAGCCAGCTCGCTTATCGTATTCAGGATTTGGG 6480



Db 6421 GTCCACAGAGAGAGGGGGCCGCAAGCAGCTCGCCTTATGCTATTCCAGATTGGGG 6480  
Qy 6481 GTTCGTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCGGTG 6540  
Db 6481 GTTCGTGTGCGAGAAATGGCCCTTTACGATGTGTCTCCACCTCCCTCAGGCGGTG 6540  
Qy 6541 ATGGGCTCTCATACGATTTCCATATCTCTCCCTGGACAGGGGTGCGAGTTCCCTGGTGAAT 6600  
Db 6541 ATGGGCTCTCTCATACGATTTCCATATCTCTCCCTGGACAGGGGTGCGAGTTCCCTGGTGAAT 6600  
Qy 6601 GCCTGGAAGCGAAGAAATGCCCTATCGGCTTCGATATGACACCCCGCTGTGTGACTCA 6660  
Db 6601 GCCTGGAAGCGAAGAAATGCCCTATCGGCTTCGATATGACACCCCGCTGTGTGACTCA 6660  
Qy 6661 ACGTCTACTGAGAATGACATCCGCTGTGAGAGTCAATCTACCAATGTGTGACTGGCC 6720  
Db 6661 ACGTCTACTGAGAATGACATCCGCTGTGAGAGTCAATCTACCAATGTGTGACTGGCC 6720  
Qy 6721 CCGGAGCCAGACAGCCATAGGTGCTCACAGAGCGGCTTTACATCGGGGGCCCGCTG 6780  
Db 6721 CCGGAGCCAGACAGCCATAGGTGCTCACAGAGCGGCTTTACATCGGGGGCCCGCTG 6780  
Qy 6781 ACTAATTTCTAAGGGCAGAACTGCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
Db 6781 ACTAATTTCTAAGGGCAGAACTGCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
Qy 6841 ACCAGCTGCGGTATACCCCTCAATGTTACTGTAAGCGCGCTGCGGCTGTGAGCTGGC 6900  
Db 6841 ACCAGCTGCGGTATACCCCTCAATGTTACTGTAAGCGCGCTGCGGCTGTGAGCTGGC 6900  
Qy 6901 AAGCTCCAGACTGACAGTCTCGTATGCGGAGACGACCTTGTGCTGTAAGTCAAGC 6960  
Db 6901 AAGCTCCAGACTGACAGTCTCGTATGCGGAGACGACCTTGTGCTGTAAGTCAAGC 6960  
Qy 6961 GCGGGGACCAAGAGGAGGAGGAGCTACGGGCTTTCAGGAGGCTATGACTAGATAC 7020  
Db 6961 GCGGGGACCAAGAGGAGGAGGAGCTACGGGCTTTCAGGAGGCTATGACTAGATAC 7020  
Qy 7021 TCTGCCCTCCCTGGGAGACCGCCCAACACAGATACGACTTGGAGTTGATCAATATGC 7080  
Db 7021 TCTGCCCTCCCTGGGAGACCGCCCAACACAGATACGACTTGGAGTTGATCAATATGC 7080  
Qy 7081 TCCTCCAAATGTGTCAGTGGCGAGTGCATCTGCAAAAGGAGTACTATCTCACCGGT 7140  
Db 7081 TCCTCCAAATGTGTCAGTGGCGAGTGCATCTGCAAAAGGAGTACTATCTCACCGGT 7140  
Qy 7141 GACCCCAACCCCTTGGCGGGCTGCGTGGGAGACAGTACACACACTCCAGTCAAT 7200  
Db 7141 GACCCCAACCCCTTGGCGGGCTGCGTGGGAGACAGTACACACACTCCAGTCAAT 7200  
Qy 7201 TCCTGGCTAGGCAACATCATGATATGCGGCCACCTTGTGGGCAAGGATGATCCTGATG 7260  
Db 7201 TCCTGGCTAGGCAACATCATGATATGCGGCCACCTTGTGGGCAAGGATGATCCTGATG 7260  
Qy 7261 ACTCATTTCTTCCATCTCTAGCTCAGGAACTTGAAGAGCCCTAGATTGTGACG 7320  
Db 7261 ACTCATTTCTTCCATCTCTAGCTCAGGAACTTGAAGAGCCCTAGATTGTGACG 7320  
Qy 7321 ATCTACGGGGCTGTACTCATTTAGCCACTTACCTACCTCAGATCAATCAACGACTC 7380  
Db 7321 ATCTACGGGGCTGTACTCATTTAGCCACTTACCTACCTCAGATCAATCAACGACTC 7380  
Qy 7381 CATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGTGAGATCAATAGGGTGGCT 7440  
Db 7381 CATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGTGAGATCAATAGGGTGGCT 7440  
Qy 7441 TCATGCTCAGGAACTTGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAGT 7500  
Db 7441 TCATGCTCAGGAACTTGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAGT 7500  
Qy 7501 GTCCGCGCTAGGCTACTGTGCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTC 7560  
Db 7501 GTCCGCGCTAGGCTACTGTGCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTC 7560

## RESULT 2

US-09-539-601-7  
; Sequence 7, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 8001  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I389/NS3-3'/wt  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1193)  
; OTHER INFORMATION: hepatitis C virus core-neomycin  
; OTHER INFORMATION: phosphotransferase fusion protein  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (1202)..(1812)  
; OTHER INFORMATION: internal ribosome entry site from  
; OTHER INFORMATION: encephalomyocarditis virus  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1813)..(7770)  
; OTHER INFORMATION: hepatitis C virus nonstructural proteins NS3-5B  
; FEATURE:  
; NAME/KEY: 3'UTR  
; LOCATION: (7771)..(8001)



1909 GTCTCCACCGCAACAATCTTTCTCTGGCGAAGCTGGTCAATGGCGTGTGGACTGTC 1968  
1921 GTCTCCACCGCAACAATCTTTCTCTGGCGAAGCTGGTCAATGGCGTGTGGACTGTC 1980  
1969 TATCATGGTCCGGGTCAAGAACCTTTGCGGCGCCAAAGGGCCCAATCAACCAATGTAC 2028  
1981 TATCATGGTCCGGGTCAAGAACCTTTGCGGCGCCAAAGGGCCCAATCAACCAATGTAC 2040  
2029 ACCAATGTGGACAGAGAACCTCGTGGCTGGCAAGCGCCCCCGGGCGGCTTCTTTGACA 2088  
2041 ACCAATGTGGACAGAGAACCTCGTGGCTGGCAAGCGCCCCCGGGCGGCTTCTTTGACA 2100  
2089 CGATGACCTGCGCGAGCTCGACACTTTTACTGGTCAAGAGGATCGGATGTCATTCG 2148  
2101 CGATGACCTGCGCGAGCTCGACACTTTTACTGGTCAAGAGGATCGGATGTCATTCG 2160  
2149 GTGCGCGCGGGGACAGAGGGGAGCTACTCTCCCAAGGCCGCTCTCTACTTTG 2208  
2161 GTGCGCGCGGGGACAGAGGGGAGCTACTCTCCCAAGGCCGCTCTCTACTTTG 2220  
2209 AAGGGCTCTTTGGGCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGCACTTTTCG 2268  
2221 AAGGGCTCTTTGGGCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGCACTTTTCG 2280  
2269 GCTGCCGTGTGACACCGAGGGGTTCGAGGGCGGTGACTTTGTACCCGTGAGTCTATG 2328  
2281 GCTGCCGTGTGACACCGAGGGGTTCGAGGGCGGTGACTTTGTACCCGTGAGTCTATG 2340  
2329 GAAACCACTATGCGGTCCCGGCTTCAACGCAAACTCTGCCCTCGGGCGGTACCGCAG 2388  
2341 GAAACCACTATGCGGTCCCGGCTTCAACGCAAACTCTGCCCTCGGGCGGTACCGCAG 2400  
2389 ACATTCAGGTGCCCCACTTACAGCCCCCTACTGGTAGCGGCAAGACATAAGGTGCCG 2448  
2401 ACATTCAGGTGCCCCACTTACAGCCCCCTACTGGTAGCGGCAAGACATAAGGTGCCG 2460  
2449 GCTGGTATGACGCCAAGGGTATAGGTGCTTGTCTGAAACCGTCCGTGCGCGCCACC 2508  
2461 GCTGGTATGACGCCAAGGGTATAGGTGCTTGTCTGAAACCGTCCGTGCGCGCCACC 2520  
2509 CTAGGTTTCGGGGGTATATGCTTAGGGACATGGTATCGACCTAACATCAGAACCGGG 2568  
2521 CTAGGTTTCGGGGGTATATGCTTAGGGACATGGTATCGACCTAACATCAGAACCGGG 2580  
2569 GTAAAGACCATCACCGGTGCCCCCTACACGTAATCCACCTATGGCAAGTTTCTTGCC 2628  
2581 GTAAAGACCATCACCGGTGCCCCCTACACGTAATCCACCTATGGCAAGTTTCTTGCC 2640  
2629 GACGGTGGTGTCTGCGGGGCGCTATGACATATATATGATGATGATGATGATGATGAT 2688  
2641 GACGGTGGTGTCTGCGGGGCGCTATGACATATATATGATGATGATGATGATGATGAT 2700  
2689 GACTCGACCACTATCTGGGCGCATCGCACAGTCTCGACCAAGCGGAGCGGTGGAGCG 2748  
2701 GACTCGACCACTATCTGGGCGCATCGCACAGTCTCGACCAAGCGGAGCGGTGGAGCG 2760  
2749 CGACTCGTGTGTGCGACCGGTACCGCTCCGGGATCGGTCAACCTATGGCAAGTTTCTTGCC 2808  
2761 CGACTCGTGTGTGCGACCGGTACCGCTCCGGGATCGGTCAACCTATGGCAAGTTTCTTGCC 2820  
2809 ATGAGGAGGTGGTCTGTGCGACCTGAGAGAAATCCCTTTTATGGCAAGGATGATGAT 2868  
2821 ATGAGGAGGTGGTCTGTGCGACCTGAGAGAAATCCCTTTTATGGCAAGGATGATGAT 2880  
2869 ATCGAGACCATCAAGGGGGGAGGACCTCATTTTCTGCAATTCACCAAGAAATGTGAT 2928  
2881 ATCGAGACCATCAAGGGGGGAGGACCTCATTTTCTGCAATTCACCAAGAAATGTGAT 2940  
2929 GAGTCCGCGGAGGCTGTGCGGCGCTCGGACTCAATGCTGATGATATTTACCGGGGCTT 2988  
2941 GAGTCCGCGGAGGCTGTGCGGCGCTCGGACTCAATGCTGATGATATTTACCGGGGCTT 3000

2989 GATGTATCGGTATACCACTAGCGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATG 3048  
3001 GATGTATCGGTATACCACTAGCGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATG 3060  
3049 ACGGGCTTTACCGGGGATTTGACTCAGTGAATCGACTGCAATACATGTGTCAACCAACA 3108  
3061 ACGGGCTTTACCGGGGATTTGACTCAGTGAATCGACTGCAATACATGTGTCAACCAACA 3120  
3109 GTGCACTTCAAGCTGAGCCGACCTTCAACATGAGACGAGCGCTGCAACGAGCGG 3168  
3121 GTGCACTTCAAGCTGAGCCGACCTTCAACATGAGACGAGCGCTGCAACGAGCGG 3180  
3169 GTGCACTTCAAGCTGAGCGGAGCGAGCTGTGTAGGGGAGAGTGGCAATTTACAGTTT 3228  
3181 GTGTACGTTCGAGCGGCGAGGAGCTGTGTAGGGGAGAGTGGCAATTTACAGTTT 3240  
3229 GTGACTTCAAGGAGAACGGGCCCTCGGSCATGTTTCAATTCCTCGGTCTGTGCGAGTCTAT 3288  
3241 GTGACTTCAAGGAGAACGGGCCCTCGGSCATGTTTCAATTCCTCGGTCTGTGCGAGTCTAT 3300  
3289 GAGCGGGCTGTGCTGTGACTACGAGCTCAGCCCGCGAGACCTCAGTTAGTTGCGGGCT 3348  
3301 GAGCGGGCTGTGCTGTGACTACGAGCTCAGCCCGCGAGACCTCAGTTAGTTGCGGGCT 3360  
3349 TACCTAAACACACCGAGGTTGCCGCTCTGCCAGACCATCTGGAGTTCTGGGAGAGCGTC 3408  
3361 TACCTAAACACACCGAGGTTGCCGCTCTGCCAGACCATCTGGAGTTCTGGGAGAGCGTC 3420  
3409 TTTTACAGGGCTCACCCACATAGACGCCCATTTCTGTCCAGACTAAGCAGGCGAGGAGAC 3468  
3421 TTTTACAGGGCTCACCCACATAGACGCCCATTTCTGTCCAGACTAAGCAGGCGAGGAGAC 3480  
3469 AACTTCCCTTACTGTGTAGCTACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCT 3528  
3481 AACTTCCCTTACTGTGTAGCTACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACCT 3540  
3529 CCATCTGGGACCAATGTGGAAGTCTCTATACGCTTAAAGCTTACGCTGACGCGGCCA 3588  
3541 CCATCTGGGACCAATGTGGAAGTCTCTATACGCTTAAAGCTTACGCTGACGCGGCCA 3600  
3589 ACCCCCTCTGTGTATAGGCTGGGAGCGGTTCAAACGAGGTTTACTACCAACACACCCCAT 3648  
3601 ACCCCCTCTGTGTATAGGCTGGGAGCGGTTCAAACGAGGTTTACTACCAACACACCCCAT 3660  
3649 ACCAAATACATATGAGTGCATGTGCGTGCAGCTGCAGTGCAGGAGCTCTGGGTG 3708  
3661 ACCAAATACATATGAGTGCATGTGCGTGCAGCTGCAGGAGCTCTGGGTG 3720  
3709 CTGTTAGGGGAGTCTTAGCAGCTCTGGCGCGTATTTGCTGACAAACAGGAGCGTGGTC 3768  
3721 CTGTTAGGGGAGTCTTAGCAGCTCTGGCGCGTATTTGCTGACAAACAGGAGCGTGGTC 3780  
3769 ATTGTGGGAGGATCATTTTGTCCGAAAGCCGCGCATCATTTCCCGACAGGGAAGTCTT 3828  
3781 ATTGTGGGAGGATCATTTTGTCCGAAAGCCGCGCATCATTTCCCGACAGGGAAGTCTT 3840  
3829 TACCGGAGGTTTGCATGAGATGGAAGTGCCTCAGCTCAGCTCCTTTACATCGAACAGGA 3888  
3841 TACCGGAGGTTTGCATGAGATGGAAGTGCCTCAGCTCAGCTCCTTTACATCGAACAGGA 3900  
3889 ATCAGCTCCCGCAACAATTTCAAACAGAGGCAATCGGGTTGCTGCAAAACAGCCACCAAG 3948  
3901 ATCAGCTCCCGCAACAATTTCAAACAGAGGCAATCGGGTTGCTGCAAAACAGCCACCAAG 3960  
3949 CAAGCGGAGGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCCCTCGAAGCCTTCTGG 4008  
3961 CAAGCGGAGGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCCCTCGAAGCCTTCTGG 4020  
4009 GCGAAGCATATGTGAATTTTCAATCAGCGGAGTACATATTTAGAGGCTTGTCCACTCTG 4068  
4021 GCGAAGCATATGTGAATTTTCAATCAGCGGAGTACATATTTAGAGGCTTGTCCACTCTG 4080  
4069 CTTGGCAACCCCGGATAGCATCACTGATGGCATTCAGGCTTATCAGCCCTTATCAGGCGGCTC 4128

Db 4081 ||||| CCTGGCAACCCCGGATAGCATCACTGATGGCAATTCAGAGCTCTATCATCAGCCCGCTC 4140  
QY 4129 ACCACCCACATACCTCTCTGTTTAAATCTCTGGGGGATGGTGGCCGCCCAACTTGCT 4188  
Db 4141 ACCACCCACATACCTCTCTGTTTAAATCTCTGGGGGATGGTGGCCGCCCAACTTGCT 4200  
QY 4189 CCTCCAGCGCTGCTCTCTCTTCTGTTAGCGCGGCATCGCTGGAGCGGCTGTGGCAGC 4248  
Db 4201 CCTCCAGCGCTGCTCTCTCTTCTGTTAGCGCGGCATCGCTGGAGCGGCTGTGGCAGC 4260  
QY 4249 ATAGGCTTGGGAAGGTGCTTGGGATATTTGGCAGGTTATGGACAGGGGTGGCAGGC 4308  
Db 4261 ATAGGCTTGGGAAGGTGCTTGGGATATTTGGCAGGTTATGGACAGGGGTGGCAGGC 4320  
QY 4309 GGGCTGTGGCCCTTTAAAGTCAATGACGGCGAGATGCCCTCAACGAGAACCTGGTTAAC 4368  
Db 4321 GGGCTGTGGCCCTTTAAAGTCAATGACGGCGAGATGCCCTCAACGAGAACCTGGTTAAC 4380  
QY 4369 CTACTCCCTGCTATCCTCTCCCTGGCGCCCTAGTGGTGGGGTGGCGCAGCGATA 4428  
Db 4381 CTACTCCCTGCTATCCTCTCCCTGGCGCCCTAGTGGTGGGGTGGCGCAGCGATA 4440  
QY 4429 CTGGCTGGGCACTGGGGCCAGGGAGGGGGCTGTGCAGTGGATGAACCGGCTGTAGCG 4488  
Db 4441 CTGGCTGGGCACTGGGGCCAGGGAGGGGGCTGTGCAGTGGATGAACCGGCTGTAGCG 4500  
QY 4489 TTGCGCTTCGGGGTAAACACGCTCTCCCCACGCACTATGCTGAGAGCGACGCTGCA 4548  
Db 4501 TTGCGCTTCGGGGTAAACACGCTCTCCCCACGCACTATGCTGAGAGCGACGCTGCA 4560  
QY 4549 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCACTGCTGAAGAGGCTTCAC 4608  
Db 4561 GCACGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGTCTGNAGAGGCTTCAC 4620  
QY 4609 CAGTGGATCAAGAGGATGTGTCCACGCCATGCTCCGGCTCGTGGCTGAAGAGATTTGG 4668  
Db 4621 CAGTGGATCAAGAGGATGTGTCCACGCCATGCTCCGGCTCGTGGCTGAAGAGATTTGG 4680  
QY 4669 GATTGGATGACAGGCTGTGACTGATTTCAAGACTGTGCTCCAGTCCAGTCCAGTCTGCGG 4728  
Db 4681 GATTGGATGACAGGCTGTGACTGATTTCAAGACTGTGCTCCAGTCCAGTCCAGTCTGCGG 4740  
QY 4729 CGATTGCCGGAGTCCCTCTCTCTCATGTCAACGTGGGTACAAGGGAGTCTGGCGGGG 4788  
Db 4741 CGATTGCCGGAGTCCCTCTCTCTCATGTCAACGTGGGTACAAGGGAGTCTGGCGGGG 4800  
QY 4789 GACGGCATATGCAAAACCACTGGCCCATGTGGAGCACAGATCAACCGGACATGTGAAAAAC 4848  
Db 4801 GACGGCATATGCAAAACCACTGGCCCATGTGGAGCACAGATCAACCGGACATGTGAAAAAC 4860  
QY 4849 GGTTCATGAGATCGTGGGCTAGGACCTGTAGTAAACGTGATGACATGGCATGGACATTCGCC 4908  
Db 4861 GGTTCATGAGATCGTGGGCTAGGACCTGTAGTAAACGTGATGACATGGCATGGACATTCGCC 4920  
QY 4909 ATTAACCGGTATACCAACCGGCGCCCTGACGCGCCCTCCCGGGGCCAAATATTCTAGGGCG 4968  
Db 4921 ATTAACCGGTATACCAACCGGCGCCCTGACGCGCCCTCCCGGGGCCAAATATTCTAGGGCG 4980  
QY 4969 CTGTGGCGGTGGCTGTGAGAGTACCTGTGAGAGTTACCGGGTGGGGATTTCCACTAC 5028  
Db 4981 CTGTGGCGGTGGCTGTGAGAGTACCTGTGAGAGTTACCGGGTGGGGATTTCCACTAC 5040  
QY 5029 GTGACGGGCATGACCACTGACAACTGAAAGTGGCCGTCTAGGTTCCGGCCCCCGAATTC 5088  
Db 5041 GTGACGGGCATGACCACTGACAACTGAAAGTGGCCGTCTAGGTTCCGGCCCCCGAATTC 5100  
QY 5089 TTCAAGAGTGGATGGGGTGGCTGTGACACAGGTACGCTCCAGCGTGAACACCCCTCCTA 5148  
Db 5101 TTCAAGAGTGGATGGGGTGGCTGTGACACAGGTACGCTCCAGCGTGAACACCCCTCCTA 5160  
QY 5149 CGGAGGAGGTACATTTCTGGTGGGGCTCAATCAATACCTGGTTGGGTCAAGCTCCCA 5208

Db 5161 CGGAGGAGGTCACTTCTCTGGTGGGGCTCAATCAATCACTGGTGGGTCAAGCTCCCA 5220  
QY 5209 TCGGAGCCCGAACCGGACGTAGAGTGTCACTTCATGCTCAACGACCCCTCCCACTT 5268  
Db 5221 TCGGAGCCCGAACCGGACGTAGAGTGTCACTTCATGCTCAACGACCCCTCCCACTT 5280  
QY 5269 AGCGCGGAGACGGCTAAAGCGTAGGTGGCCAGGGATCTCCCCCTCTTGGCCAGCTCA 5328  
Db 5281 AGCGCGGAGACGGCTAAAGCGTAGGTGGCCAGGGATCTCCCCCTCTTGGCCAGCTCA 5340  
QY 5329 TCAGCTAGCAGCTGTCTGGCTCTCTTTGAAGGCAACATGCACTACCCGTCACTACTCC 5388  
Db 5341 TCAGCTAGCAGCTGTCTGGCTCTCTTTGAAGGCAACATGCACTACCCGTCACTACTCC 5400  
QY 5389 CGGACGCTGACCTCATGAGGCCAACCTCTCTGTGGCGGACGAGATGGCGGGAACTC 5448  
Db 5401 CGGACGCTGACCTCATGAGGCCAACCTCTCTGTGGCGGACGAGATGGCGGGAACTC 5460  
QY 5449 ACCCGCTGGAGTCAAGAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAGCG 5508  
Db 5461 ACCCGCTGGAGTCAAGAAATAAGGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCAAGCG 5520  
QY 5509 GAGGAGGATGAGAGGAAAGTATCCCTTCGGCGGAGATCTCTCGGAGGTCCAGGAAATTC 5568  
Db 5521 GAGGAGGATGAGAGGAAAGTATCCCTTCGGCGGAGATCTCTCGGAGGTCCAGGAAATTC 5580  
QY 5569 CCTCAGCGATGCCCATATGGGCAACCGCGGANTTAAACCCCTCCACTGTTAGAGTCTCGG 5628  
Db 5581 CCTCAGCGATGCCCATATGGGCAACCGCGGANTTAAACCCCTCCACTGTTAGAGTCTCGG 5640  
QY 5629 AAGGACCCGGACTAGTCCCTCCAGTGTGACAGGGTGTCCATTGCGCGCTGCCAGGCG 5688  
Db 5641 AAGGACCCGGACTAGTCCCTCCAGTGTGACAGGGTGTCCATTGCGCGCTGCCAGGCG 5700  
QY 5689 CCTCGATACCACTCCACGGAGGAGGAGCGGTTGTCTGTGACAGATCTACCGTGTCT 5748  
Db 5701 CCTCGATACCACTCCACGGAGGAGGAGCGGTTGTCTGTGACAGATCTACCGTGTCT 5760  
QY 5749 TCTGCTTGGCGGAGCTCGCCAAAGACCTTCGGCAGCTCGAATCTCGGCGCTCGAC 5808  
Db 5761 TCTGCTTGGCGGAGCTCGCCAAAGACCTTCGGCAGCTCGAATCTCGGCGCTCGAC 5820  
QY 5809 AGCGGACCGCAACCGGCTCTCTGACGAGCCCTCCGACGAGCGGACGCGGGAATCCGAC 5868  
Db 5821 AGCGGACCGCAACCGGCTCTCTGACGAGCCCTCCGACGAGCGGACGCGGGAATCCGAC 5880  
QY 5869 GTTGAGTGTACTCTCTCATGCCCCCTTGGAGGGGAGCGGGGATCCCGATCTCAGC 5928  
Db 5881 GTTGAGTGTACTCTCTCATGCCCCCTTGGAGGGGAGCGGGGATCCCGATCTCAGC 5940  
QY 5929 GACGGTCTTGGTCTACCGTAAAGCGAGGAGCTGTAGGAGCTGTGTGTCTGATG 5988  
Db 5941 GACGGTCTTGGTCTACCGTAAAGCGAGGAGCTGTAGGAGCTGTGTGTCTGATG 6000  
QY 5989 TCCTACATGACGACGCGCTGATACGCCATGCGCTCGGAGGAACCAAGTGGCC 6048  
Db 6001 TCCTACATGACGACGCGCTGATACGCCATGCGCTCGGAGGAACCAAGTGGCC 6060  
QY 6049 ATCAATGACGTGAGCACTCTTGTCTCGTCACCAACCTTGGTCTATGTCTACACATCT 6108  
Db 6061 ATCAATGACGTGAGCACTCTTGTCTCGTCACCAACCTTGGTCTATGTCTACACATCT 6120  
QY 6109 CGGAGCGAAGCTTGGCGCAGAAAGGTACCTTTGACAGACTGAGGTCCTGGACGAC 6168  
Db 6121 CGGAGCGAAGCTTGGCGCAGAAAGGTACCTTTGACAGACTGAGGTCCTGGACGAC 6180  
QY 6169 CACTACCGGACGTGTCTCAAGGAGATGAAGCGAAGGCTCCACAGTTAAGCTTAACTT 6228  
Db 6181 CACTACCGGACGTGTCTCAAGGAGATGAAGCGAAGGCTCCACAGTTAAGCTTAACTT 6240  
QY 6229 CTATCCGTGGAGGAGCTGTAAAGTGAAGCCCCCAATTCGGCCAGATCTAAATTTGGC 6288  
Db 6241 CTATCCGTGGAGGAGCTGTAAAGTGAAGCCCCCAATTCGGCCAGATCTAAATTTGGC 6300

QY 6289 TATGGGCAAGAGCGTCCGGAACCTATCCAGCAAGCCGTTAACCACATCCGCTCCGTG 6348  
DB 6301 TATGGGCAAGAGAGCTCCGGAACCTATCCAGCAAGCCGTTAACCACATCCGCTCCGTG 6360  
QY 6349 TGGAGGACTTGTCTGGAAGACACTGAGACACCAATTGACACACCATCATGGCAAAAAT 6408  
DB 6361 TGGAGGACTTGTCTGGAAGACACTGAGACACCAATTGACACACCATCATGGCAAAAAT 6420  
QY 6409 GAGGTTTCTCGCTCCAAACAGAGAGAGGGGGCCGCAAGCCAGCTCGCTTTATCGTATTC 6468  
DB 6421 GAGGTTTCTCGCTCCAAACAGAGAGAGGGGGCCGCAAGCCAGCTCGCTTTATCGTATTC 6480  
QY 6469 CCAGATTGGGGGTTTCGTGTCGGAGAAATGCCCCCTTTACGATGTCCTCCACCCCTC 6528  
DB 6481 CCAGATTGGGGGTTTCGTGTCGGAGAAATGCCCCCTTTACGATGTCCTCCACCCCTC 6540  
QY 6529 CCTCAGGCGGTGATGGGCTCTTCATACGGAATTCGAATCTCTCTGGACAGCGGTCCAG 6588  
DB 6541 CCTCAGGCGGTGATGGGCTCTTCATACGGAATTCGAATCTCTCTGGACAGCGGTCCAG 6600  
QY 6589 TTCCTGTTGAATGCTTGGAAAGCAAGAAATGCCCCCTTTACGATGTCCTCCACCCCTC 6648  
DB 6601 TTCCTGTTGAATGCTTGGAAAGCAAGAAATGCCCCCTTTACGATGTCCTCCACCCCTC 6660  
QY 6649 TGTCTTCACTCAAGGTCACATGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATGT 6708  
DB 6661 TGTCTTCACTCAAGGTCACATGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATGT 6720  
QY 6709 TGTGACTTGGCCCCCGGAGCAGACAGGCCATAAGGTCGCTTCACAGAGCGGCTTTTACATC 6768  
DB 6721 TGTGACTTGGCCCCCGGAGCAGACAGGCCATAAGGTCGCTTCACAGAGCGGCTTTTACATC 6780  
QY 6769 GGGGGCCCCCTGACTAATCTTAAAGGCAAGAAATGCCCCCTTTACGATGTCCTCCACCCCTC 6828  
DB 6781 GGGGGCCCCCTGACTAATCTTAAAGGCAAGAAATGCCCCCTTTACGATGTCCTCCACCCCTC 6840  
QY 6829 GGTCTACTGACGACGCTGGGTAAATACCTCTACATGTTACTTGAAGCGCGCTCGCGCC 6888  
DB 6841 GGTCTACTGACGACGCTGGGTAAATACCTCTACATGTTACTTGAAGCGCGCTCGCGCC 6900  
QY 6889 TGTGAGCTGGGAAGCTCCAGGACTGACGAGTGTCTGATGTCGAGACGACCTTGTGCTT 6948  
DB 6901 TGTGAGCTGGGAAGCTCCAGGACTGACGAGTGTCTGATGTCGAGACGACCTTGTGCTT 6960  
QY 6949 ATCTGTGAAGCGGGGAGCCAAAGAGACGAGCGAGCCCTTACGGGCTTCCAGGAGCT 7008  
DB 6961 ATCTGTGAAGCGGGGAGCCAAAGAGACGAGCGAGCCCTTACGGGCTTCCAGGAGCT 7020  
QY 7009 ATGACTAGTACTTGTGCCCCCTTGGGACCGCGCCCAACCCAGAAATAGGAGTTG 7068  
DB 7021 ATGACTAGTACTTGTGCCCCCTTGGGACCGCGCCCAACCCAGAAATAGGAGTTG 7080  
QY 7069 ATAACTATGCTCTCTCAATGTCTCAGTCTGCGCACGATGATCTGCAAAAGGTTGATC 7128  
DB 7081 ATAACTATGCTCTCTCAATGTCTCAGTCTGCGCACGATGATCTGCAAAAGGTTGATC 7140  
QY 7129 TATCTACCGGTGACCCCAACACCCCTTGTGCGGGCTGCGTGGGAGACAGCTTAGACAC 7188  
DB 7141 TATCTACCGGTGACCCCAACACCCCTTGTGCGGGCTGCGTGGGAGACAGCTTAGACAC 7200  
QY 7189 ACTCCAGTCAATCTCTGGCTAGGCAACATCATCATGATGATGCGGCCACCTTGTGGGCAAG 7248  
DB 7201 ACTCCAGTCAATCTCTGGCTAGGCAACATCATCATGATGATGCGGCCACCTTGTGGGCAAG 7260  
QY 7249 ATGATCTGATGACTCTTTCTTCTCTCAATCTCTCTAGCTCAGGAACAACCTTGAAGAGCC 7308  
DB 7261 ATGATCTGATGACTCTTTCTTCTCTCAATCTCTCTAGCTCAGGAACAACCTTGAAGAGCC 7320  
QY 7309 CTAGATTGTCAAGTCTACGGGGCTGTTTACTTCCATTGAGCCACTTACCTCAGATC 7368  
DB 7321 CTAGATTGTCAAGTCTACGGGGCTGTTTACTTCCATTGAGCCACTTACCTCAGATC 7380

QY 7369 ATTCAACGACTCCATGGCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATC 7428  
DB 7381 ATTCAACGACTCCATGGCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATC 7440  
QY 7429 AATAGGCTGGCTTTCATCGCTCAGGAAACTTGGGGTACCGCCCTTTCGAGTCTCGAGACAT 7488  
DB 7441 AATAGGCTGGCTTTCATCGCTCAGGAAACTTGGGGTACCGCCCTTTCGAGTCTCGAGACAT 7500  
QY 7489 CGGCCCAGAGTGTTCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCACTTGTGGC 7548  
DB 7501 CGGCCCAGAGTGTTCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCACTTGTGGC 7560  
QY 7549 AAGTACCTCTTCAACTGGGAGTAAGACCAAGCTCAAACTCACTCAATCCCGGCTGG 7608  
DB 7561 AAGTACCTCTTCAACTGGGAGTAAGACCAAGCTCAAACTCACTCAATCCCGGCTGG 7620  
QY 7609 TCCAGTGTGATTTTATCAGCTGTTGCTTGTGTTACAGCGGGGAGACATATATCAC 7668  
DB 7621 TCCAGTGTGATTTTATCAGCTGTTGCTTGTGTTACAGCGGGGAGACATATATCAC 7680  
QY 7669 AGCTGTCTGTGCGCGACCCCGCTGTTCAATGTGTTGCTTACTCTCTTCTGTAGGG 7728  
DB 7681 AGCTGTCTGTGCGCGACCCCGCTGTTCAATGTGTTGCTTACTCTCTTCTGTAGGG 7740  
QY 7729 GTAGGCATCTATCTACTCCCAACCGGATGACCGGGACCTAAACACTCCAGGCCAATAGG 7788  
DB 7741 GTAGGCATCTATCTACTCCCAACCGGATGACCGGGAGCTAAACACTCCAGGCCAATAGG 7800  
QY 7789 CCATCTCTGTTTTTCCCTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7848  
DB 7801 CCATCTCTGTTTTTCCCTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860  
QY 7849 TTTTCTCTTTTTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTAGC 7908  
DB 7861 TTTTCTCTTTTTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTTTTCTTTAGC 7920  
QY 7909 CCTAGTCAAGCTAGCTGTGAAGGTCCGTGAGCGCTTGAATGAGTGTGCTGAGTGTGATC 7968  
DB 7921 CCTAGTCAAGCTAGCTGTGAAGGTCCGTGAGCGCTTGAATGAGTGTGCTGAGTGTGATC 7980  
QY 7969 TGGCTCTCTGACATCAAGT 7989  
DB 7981 TGGCTCTCTGACATCAAGT 8001

RESULT 3  
US-09-539-601-22  
; Sequence 22, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 22  
; LENGTH: 8001  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I389/NS3-3'/5.1  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1193)  
; OTHER INFORMATION: hepatitis C virus core - neomycin  
; OTHER INFORMATION: phosphotransferase fusion protein  
; FEATURE:

```

; NAME/KEY: RBS
; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(7770)
; OTHER INFORMATION: hepatitis C virus nonstructural proteins NS3 - 5B
; OTHER INFORMATION: of cell culture-adapted clone no. 5.1
; FEATURE:
; NAME/KEY: 5'UTR
; LOCATION: (7771)..(8001)
; US-09-539-601-22

Query Match          99.5%; Score 7947.8; DB 4; Length 8001;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 7977; Conservative 0; Mismatches 12; Indels 12; Gaps 1;

QY 1 GCAGCCCCCGATTGGGGGGGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTG 60
Db 1 GCAGCCCCCGATTGGGGGGGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTG 60

QY 61 TCTTCAAGCAGAAAGCGTCTAGCCATGGGTTAGTATGATGTCTGTGAGCCCTCCAGGAC 120
Db 61 TCTTCAAGCAGAAAGCGTCTAGCCATGGGTTAGTATGATGTCTGTGAGCCCTCCAGGAC 120

QY 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGAG 180
Db 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATTCGCGAG 180

QY 181 GACGACCGGGTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTGGGGGTGCCCCC 240
Db 181 GACGACCGGGTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTGGGGGTGCCCCC 240

QY 241 GCGAGACTCTAGCCAGTAGTGTCTGGTCCGAAAGGCGTGTGCTACTGCTGTATG 300
Db 241 GCGAGACTCTAGCCAGTAGTGTCTGGTCCGAAAGGCGTGTGCTACTGCTGTATG 300

QY 301 GTGCTTGCAGTGCCTCCGGGAGTCTCGTAGACCGTGCACCATGAGCAGCAATCCATAAC 360
Db 301 GTGCTTGCAGTGCCTCCGGGAGTCTCGTAGACCGTGCACCATGAGCAGCAATCCATAAC 360

QY 361 CTCAAGAAAAAACCCAAA-----GGGCGCGCCATGATTGAACAAGATTGGATTGC 408
Db 361 CTCAAGAAAAAACCCAAA-----GGGCGCGCCATGATTGAACAAGATTGGATTGC 420

QY 409 AGCAGGTTCTCGGCGGCTTGGGTGAGAGGCTATTGCTATGATCGGACACACAGA 468
Db 409 AGCAGGTTCTCGGCGGCTTGGGTGAGAGGCTATTGCTATGATCGGACACACAGA 480

QY 469 CAATCGGCTGCTCTGATGCGCGTGTTCGGGCTGTGAGCGAGGGGCGCGGTTCTTT 528
Db 469 CAATCGGCTGCTCTGATGCGCGTGTTCGGGCTGTGAGCGAGGGGCGCGGTTCTTT 540

QY 529 TTGTCAAGACGACCTGTCCGTGCGCTTGAATGAATGAGGACGAGGACGCGGCTAT 588
Db 529 TTGTCAAGACGACCTGTCCGTGCGCTTGAATGAATGAGGACGAGGACGCGGCTAT 600

QY 589 CGTGGCTGCGCAGGAGGCGCTTCTTGGCAGCTGTGCTCGACGTTGTCACTGAAGCGG 648
Db 589 CGTGGCTGCGCAGGAGGCGCTTCTTGGCAGCTGTGCTCGACGTTGTCACTGAAGCGG 660

QY 649 GAAGGACTGGGTGCTGCTATTGGGGGAGAGTCCGGGGGAGGATCTCCTGTCACTGCTTG 708
Db 649 GAAGGACTGGGTGCTGCTATTGGGGGAGAGTCCGGGGGAGGATCTCCTGTCACTGCTTG 720

QY 709 CTCTTCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATPACGCTTGATC 768
Db 709 CTCTTCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATPACGCTTGATC 780

QY 769 CGGCTACCTGCCATTCGACCAACGAGCAATCATCGATCGAGCGGACGCTACTTCGGA 828
Db 769 CGGCTACCTGCCATTCGACCAACGAGCAATCATCGATCGAGCGGACGCTACTTCGGA 840
```



1909 GTCTCCACCGCAACACAACTTTCTGCGGACCTCGCTCAATGCGGTGTTGGACTGTC 1968  
1921 GTCTCCACCGCAACAACTTTCTGCGGACCTCGCTCAATGCGGTGTTGGACTGTC 1980  
1969 TATCATGGTGCCTGCTCAAGACCTTTGCGGCGCCAAAGGCGCCCAATCACCCAAATGTAC 2028  
1981 TATCATGGTGCCTGCTCAAGACCTTTGCGGCGCCAAAGGCGCCCAATCACCCAAATGTAC 2040  
2029 ACCAATGTGACACGAGGACCTGCTGCGGTGGAAGGCGCCCGGCGCGGTTCCTTGACA 2088  
2041 ACCAATGTGACACGAGGACCTGCTGCGGTGGAAGGCGCCCGGCGCGGTTCCTTGACA 2100  
2089 CCATGCACTGCGGACGCTCGGACCTTTACTTGTGTACAGAGCATGCCGATGCTATCCG 2148  
2101 CCATGCACTGCGGACGCTCGGACCTTTACTTGTGTACAGAGCATGCCGATGCTATCCG 2160  
2149 GTGCGCGCGGCGGCGACAGCAGAGGCGAGCCTACTCTCCCGCAGGCGCGCTCTCCTACTTG 2208  
2161 GTGCGCGCGGCGGCGACAGCAGGCGGAGCCTACTCTCCCGCAGGCGCGCTCTCCTACTTG 2220  
2209 AAGGGCTCTTGGCGGCTGCTGCTGCTGCGGCGGCGAGCGCTGTGGGCGATCTTTCGG 2268  
2221 AAGGGCTCTTGGCGGCTGCTGCTGCTGCGGCGGCGAGCGCTGTGGGCGATCTTTCGG 2280  
2269 GCTGCCGTGTGCACCCGAGGGGTTTGGAAAGCGGTGGACTTTGTACCGCTCGAGTCTATG 2328  
2281 GCTGCCGTGTGCACCCGAGGGGTTTGGAAAGCGGTGGACTTTGTACCGCTCGAGTCTATG 2340  
2329 GAAACCACTATGCGGTCCCCGGTCTTCAAGCAAACTCGTCCCCCGGCGGTACCGCAG 2388  
2341 GAAACCACTATGCGGTCCCCGGTCTTCAAGCAAACTCGTCCCCCGGCGGTACCGCAG 2400  
2389 ACATTCAGGTGCGCCATCTACAGCGCCCTACTGGTAGCGCAAGAGCACTAAGGTGCG 2448  
2401 ACATTCAGGTGCGCCATCTACAGCGCCCTACTGGTAGCGCAAGAGCACTAAGGTGCG 2460  
2449 GCTGCGTATGCAAGGCTTAAGGTGCTGCTGTAACCGCTCGCTGCGCGCGCAC 2508  
2461 GCTGCGTATGCAAGGCTTAAGGTGCTGCTGTAACCGCTCGCTGCGCGCGCAC 2520  
2509 CTAGGTTTCGGGCGGTATGCTTAAGGCAATGATGATCGACCTTAACATCAGAACCGGG 2568  
2521 CTAGGTTTCGGGCGGTATGCTTAAGGCAATGATGATCGACCTTAACATCAGAACCGGG 2580  
2569 GTAAGGACCATCACAGCGGTGCGCCATCAGTACTCGACCTATGGAAGTTCTTGCC 2628  
2581 GTAAGGACCATCACAGCGGTGCGCCATCAGTACTCGACCTATGGAAGTTCTTGCC 2640  
2629 GACGTTGTTGCTCTGGGGCGCTTATGACATCATATATGATGATGATGATGATGATGAT 2688  
2641 GACGTTGTTGCTCTGGGGCGCTTATGACATCATATATGATGATGATGATGATGATGAT 2700  
2689 GACTCGACCATCTCTGGGCGATCGGCACAGTCTCGACCAAGCGGAGAGCGGTGGAGCG 2748  
2701 GACTCGACCATCTCTGGGCGATCGGCACAGTCTCGACCAAGCGGAGAGCGGTGGAGCG 2760  
2749 CGACTCGTCTGCTCGGCGACCGCTACCGCTCGGGATCGGTACCGGTGCGACATCCAAAC 2808  
2761 CGACTCGTCTGCTCGGCGACCGCTACCGCTCGGGATCGGTACCGGTGCGACATCCAAAC 2820  
2809 ATCGAGAGGTGCTGCTGCTCAGCACTGGAGAAATCCCTTTTATGCAAGGCGATCCCC 2868  
2821 ATCGAGAGGTGCTGCTGCTCAGCACTGGAGAAATCCCTTTTATGCAAGGCGATCCCC 2880  
2869 ATCGAGACCATCAAGGGGGGAGGACCTCATTTTCTGCCATTCCAAAGAAAGATGTGAT 2928  
2881 ATCGAGACCATCAAGGGGGGAGGACCTCATTTTCTGCCATTCCAAAGAAAGATGTGAT 2940  
2929 GAGTCCGCGGAGAGCTGTGCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCGCTT 2988  
2941 GAGTCCGCGGAGAGCTGTGCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCGCTT 3000  
2989 GATGTATCCGTATACCAACTAGCGGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATG 3048

3001 GATGTATCCGTATACCAACTAGCGGAGAGCTCATTTGCTGTAGCAACGAGCGCTCTAATG 3060  
3049 ACGGGCTTTACCGCGGATTTTGCATCTAGTGTATCTGACTGCAATACATGTGTCAACAGACA 3108  
3061 ACGGGCTTTACCGGTGATTTGCACTCAGTGTATCGACTGCAATACATGTGTCAACAGACA 3120  
3109 GTGCACTTACGCTTGAGACCGGACCTTCACTTACGAGACGAGCGGTGCCACAGACGCG 3168  
3121 GTGCACTTACGCTTGAGACCGGACCTTCACTTACGAGACGAGCGGTGCCACAGACGCG 3180  
3169 GTGCACTTACGCTTGAGACCGGACCTTCACTTACGAGACGAGCGGTGCCACAGACGCG 3228  
3181 GTGCACTTACGCTTGAGACCGGACCTTCACTTACGAGACGAGCGGTGCCACAGGTTT 3240  
3229 GTGCACTTACGAGAAACCGGCGCTCGGGCATGTTTCATTTCTCGGTTCTGTGCGAGTGTAT 3288  
3241 GTGACTCCAGGAGAACCGGCGCTCGGGCATGTTTCATTTCTCGGTTCTGTGCGAGTGTAT 3300  
3289 GACCGGGCTGTGCTTGTGTACGAGCTCAGCGCGCGCGAGACCTCAGTTAGTTGCGGGCT 3348  
3301 GACCGGGCTGTGCTTGTGTACGAGCTCAGCGCGCGCGAGACCTCAGTTAGTTGCGGGCT 3360  
3349 TACCTAAACACACAGGCTTGGCGCTCTGCCAGAACCATCTGGAGTTCTGGGAGAGGTC 3408  
3361 TACCTAAACACACAGGCTTGGCGCTCTGCCAGAACCATCTGGAGTTCTGGGAGAGGTC 3420  
3409 TTTACAGGCGCTCACCCACATAGACGCGCATTTCTTTGTCCCAGACTAAGCAGGAGGAGAC 3468  
3421 TTTACAGGCGCTCACCCACATAGACGCGCATTTCTTTGTCCCAGACTAAGCAGGAGGAGAC 3480  
3469 AACTTCCCTTACCTGTGTAGCATACCGGCTACGGTGTGGCGCAGGGCTCAGGCTCCACCT 3528  
3481 AACTTCCCTTACCTGTGTAGCATACCGGCTACGGTGTGGCGCAGGGCTCAGGCTCCACCT 3540  
3529 CCATCTGGGACCAAAATGTGGAAGTCTCATACGGCTAAAGCTACGCTGCAAGGCGCA 3588  
3541 CCATCTGGGACCAAAATGTGGAAGTCTCATACGGCTAAAGCTACGCTGCAAGGCGCA 3600  
3589 ACGCCCTGTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACACCCATA 3648  
3601 ACGCCCTGTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACCCATA 3660  
3649 ACCAATACATCATGGAATGCAATGCGCTGACCTGTGAGTGTGTCACAGACCTCGGGT 3708  
3661 ACCAATACATCATGGAATGCAATGCGCTGACCTGTGAGTGTGTCACAGACCTCGGGT 3720  
3709 CTGCTAGGCGGAGTCTTACGAGCTCTGCGCGCTGATTCCTGCAACAGCAGCGTGGTC 3768  
3721 CTGCTAGGCGGAGTCTTACGAGCTCTGCGCGCTGATTCCTGCAACAGCAGCGTGGTC 3780  
3769 ATTGTGGGAGGATCATTTGTCCGAAAGCGCGCATCATTTCCGACAGGAGTCCCT 3828  
3781 ATTGTGGGAGGATCATTTGTCCGAAAGCGCGCATCATTTCCGACAGGAGTCCCT 3840  
3829 TACCGGAGTTTCATGAGATGGAAGTGGCGCTCACACCTCCCTTATACATCGAACAGGGA 3888  
3841 TACCGGAGTTTCATGAGATGGAAGTGGCGCTCACACCTCCCTTATACATCGAACAGGGA 3900  
3889 ATGAGCTCGCGCAACAAATTCAAACGAAGGCAATCGGGTGTGCAACAGCAGCCCAAG 3948  
3901 ATGAGCTCGCGCAACAAATTCAAACGAAGGCAATCGGGTGTGCAACAGCAGCCCAAG 3960  
3949 CAAGCGGAGGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCCCTCGAAGGCTTCTGG 4008  
3961 CAAGCGGAGGCTGCTGCTCCCGTGGTGGAAATCAAGTGGCGGACCCCTCGAAGGCTTCTGG 4020  
4009 GCGAAGCATATGTGGAATTTTCATCAGCGGATCAATATTTAGAGGCTTGTCCACTCTG 4068  
4021 GCGAAGCATATGTGGAATTTTCATCAGCGGATCAATATTTAGAGGCTTGTCCACTCTG 4080  
4069 CTTGGCAACCCCGGATAGCATCAGTGGCAATTCAGGCTTATCAGGCTTATCAGGCGGCTC 4128



6289 TATGGGCAAGACGCTCGGAAACCTATCCAGAAAGCCGCTTAACACATCCCGTCG 6348  
6301 TATGGGCAAGACGCTCGGAAACCTATCCAGAAAGCCGCTTAACACATCCCGTCG 6360  
6349 TGAAGGACTTGTGGAAGACACTGAGACACCAATTGACACCAACCATCATGGCAAAAT 6408  
6361 TGAAGGACTTGTGGAAGACACTGAGACACCAATTGACACCAACCATCATGGCAAAAT 6420  
6409 GAGGTTTTCGCTCCAAACAGAGAAAGGGGGGGGCGAAGCAGTTCGCTTATCGTATTC 6468  
6421 GAGGTTTTCGCTCCAAACAGAGAAAGGGGGGGGCGAAGCAGTTCGCTTATCGTATTC 6480  
6469 CAGATTGTTGGGGTTCGTGTGTCGAGAAATGSCCTTTACGATGTGCTCCACCTC 6528  
6481 CAGATTGTTGGGGTTCGTGTGTCGAGAAATGSCCTTTACGATGTGCTCCACCTC 6540  
6529 CCTCAGGCGGTGATGGCTCTTCCATCGGATTCCTCCTGACAGCGGTCGAG 6588  
6541 CCTCAGGCGGTGATGGCTCTTCCATCGGATTCCTCCTGACAGCGGTCGAG 6600  
6589 TTTCCTGTAATGCTGGAAGGAGAAATGCTTATGGCTTCGATATGACACCGC 6648  
6601 TTTCCTGTAATGCTGGAAGGAGAAATGCTTATGGCTTCGATATGACACCGC 6660  
6649 TGTGTTGACTCAACGCTCACTGAGATGACATCGCTGTGAGGAGTCAATCTACCAATG 6708  
6661 TGTGTTGACTCAACGCTCACTGAGATGACATCGCTGTGAGGAGTCAATCTACCAATG 6720  
6709 TGTGACTTGGCCCCGGAAGCAGACAGGCCATTAAGTTCGCTCAGAGCGGCTTTATATC 6768  
6721 TGTGACTTGGCCCCGGAAGCAGACAGGCCATTAAGTTCGCTCAGAGCGGCTTTATATC 6780  
6769 GGGGGCCCCCTGACTTAATCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGGAGC 6828  
6781 GGGGGCCCCCTGACTTAATCTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGGAGC 6840  
6829 GGTGTACTGACGACACGCTCGGTGTAATACCTCACAATGTTACTTGAAGGCGCTGCGGCC 6888  
6841 GGTGTACTGACGACACGCTCGGTGTAATACCTCACAATGTTACTTGAAGGCGCTGCGGCC 6900  
6889 TGTGAGCTCGAAGCTTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTGCTT 6948  
6901 TGTGAGCTCGAAGCTTCCAGGACTGACGATGCTCGTATGCGGAGACGACCTTGTGCTT 6960  
6949 ATCTGTGAAGCGGGGACCCAAAGAGAGAGAGCGGAGCCCTACGGGCTTTCACGAGGCT 7008  
6961 ATCTGTGAAGCGGGGACCCAAAGAGAGAGAGCGGAGCCCTACGGGCTTTCACGAGGCT 7020  
7009 ATGACTAGATACTCTGCCCCCTCGGGGACCCGCAAAACAGAAATACGACTTGGAGTTG 7068  
7021 ATGACTAGATACTCTGCCCCCTCGGGGACCCGCAAAACAGAAATACGACTTGGAGTTG 7080  
7069 ATAACATCATGCTCCTCCAATGTTGCTAGTTCGCGCAGCATGCTGTCGCAAAAGGCTGAC 7128  
7081 ATAACATCATGCTCCTCCAATGTTGCTAGTTCGCGCAGCATGCTGTCGCAAAAGGCTGAC 7140  
7129 TATCTACCGGTGACCCCAACACCCCTTGGCGGGCTGCGGTGGGAGACAGCTAGACAC 7188  
7141 TATCTACCGGTGACCCCAACACCCCTTGGCGGGCTGCGGTGGGAGACAGCTAGACAC 7200  
7189 ACTCCAGTCAATCTCGCTGAGCAACATCATATGATGTCGCCACCTTGTGGGCAAGG 7248  
7201 ACTCCAGTCAATCTCGCTGAGCAACATCATATGATGTCGCCACCTTGTGGGCAAGG 7260  
7249 ATGATCTGATGACTCAATTTCTTCCATCCTTCTAGCTCAGGAACCACTTGAAGAAAGCC 7308  
7261 ATGATCTGATGACTCAATTTCTTCCATCCTTCTAGCTCAGGAACCACTTGAAGAAAGCC 7320  
7309 CTGATTTGTGAGTCTACGGGCGCTGTTACTCCATTGAGGACCTTACCTCAGATC 7368  
7321 CTGATTTGTGAGTCTACGGGCGCTGTTACTCCATTGAGGACCTTACCTCAGATC 7380  
7369 ATTCAAGACTCCATGSCCTTAGGCGATTTTCACTCCATAGTACTCTCCAGGTGAGATC 7428

7381 ATTCAAGACTCCATGGCCTTAGCGCATTTTCACTCCATAGTTACTCTCAGGTGAGATC 7440  
7429 AATAGGTTGGCTTTCATGCTCCTCAGAAACTTGGGTTACCGCCTTGGAGTCTGGAGACAT 7488  
7441 AATAGGTTGGCTTTCATGCTCCTCAGAAACTTGGGTTACCGCCTTGGAGTCTGGAGACAT 7500  
7489 CGGGCCAGAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGGAGGCTGCACTTGTGGC 7548  
7501 CGGGCCAGAAGTGTCCGCGCTAGGCTACTGTCCAGGGGGGGAGGCTGCACTTGTGGC 7560  
7549 AAGTACTCTTCAACTTGGGAGTAAGCAACAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7608  
7561 AAGTACTCTTCAACTTGGGAGTAAGCAACAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7620  
7609 TCCAGTTGGAATTTATCCAGCTGGTTGCTGTTTACAGCGGGGAGACATATATCAC 7668  
7621 TCCAGTTGGAATTTATCCAGCTGGTTGCTGTTTACAGCGGGGAGACATATATCAC 7680  
7669 AGCTGCTGTCGTCGCGACCCCGCTGGTTCACTGTTGCTGCTACTCCTTCTCTAGGG 7728  
7681 AGCTGCTGTCGTCGCGACCCCGCTGGTTCACTGTTGCTGCTACTCCTTCTCTAGGG 7740  
7729 GTAGGCACTATCTACTTCCCAACCGATGAACCGGACCTTAAACACTCCAGGCAATAGG 7788  
7741 GTAGGCACTATCTACTTCCCAACCGATGAACCGGAGCTTAAACACTCCAGGCAATAGG 7800  
7789 CCAATCTGTTTTTTTCCCTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7848  
7801 CCAATCTGTTTTTTTCCCTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860  
7849 TTTTCTCTTTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTAGC 7908  
7861 TTTTCTCTTTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTTTTTTCTTTAGC 7920  
7909 CCTAGTCAAGGCTAGCTGTGAAAGGTCGCTGAGCGCTTGACTGACAGAGTGTGTATAC 7968  
7921 CCTAGTCAAGGCTAGCTGTGAAAGGTCGCTGAGCGCTTGACTGACAGAGTGTGTATAC 7980  
7969 TGGCCTCTCTGAGATCAAGT 7989  
7981 TGGCCTCTCTGAGATCAAGT 8001

RESULT 4  
US-09-539-601-16  
; Sequence 16, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 16  
; LENGTH: 8001  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I389/NS3-3'/9-13F  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1193)  
; OTHER INFORMATION: hepatitis C virus core-neomycin phosphotransferase  
; FEATURE:  
; NAME/KEY: RBS

```

; LOCATION: (1202)...(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)...(7770)
; OTHER INFORMATION: hepatitis C virus nonstructural protein NS3-5B;
; OTHER INFORMATION: carries cell culture-adaptive mutations from clone
; OTHER INFORMATION: 9-13F
; FEATURE:
; NAME/KEY: 3' UTR
; LOCATION: (7771)...(8001)
; US-09-539-601-16

Query Match          99.5%; Score 7946.2; DB 4; Length 8001;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 7976; Conservative 0; Mismatches 13; Indels 12; Gaps 1;

QY 1 GCCAGCCCCGATTTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
Dd |||
1 GCCAGCCCCGATTTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
QY 61 TCTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTCGAGCCTCCAGGAC 120
Dd |||
61 TCTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTATGATGTCGTCGAGCCTCCAGGAC 120
QY 121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGGACCGGTGAGTACACCGGAATGCCAG 180
Dd |||
121 CCCCCCTCCGGGAGAGCCATAGTGTCTGCGGACCGGTGAGTACACCGGAATGCCAG 180
QY 181 GACGACCGGGTCTTCTTTGGATCAACCCGCTCAATGCTCGGAGATTTGGGGTGCCCC 240
Dd |||
181 GACGACCGGGTCTTCTTTGGATCAACCCGCTCAATGCTCGGAGATTTGGGGTGCCCC 240
QY 241 GCGAGCTCTAGCCGAGTAGTGTGGTTCGCGAAAGCGTCTGTCGTCGTCGTCGTCGTC 300
Dd |||
241 GCGAGCTCTAGCCGAGTAGTGTGGTTCGCGAAAGCGTCTGTCGTCGTCGTCGTCGTC 300
QY 301 GTGCTTGGAGTCCCGGGAGTCTGCTGAGCCGTCACATGAGCAGCAATCCGTAAC 360
Dd |||
301 GTGCTTGGAGTCCCGGGAGTCTGCTGAGCCGTCACATGAGCAGCAATCCGTAAC 360
QY 361 CTCGAAGAAAAACCAAA-----GGGCGCGCCATGATGAACAAGATGGATTGC 408
Dd |||
361 CTCGAAGAAAAACCAAAACGTAACCAACCGCGCGCCATGATGAACAAGATGGATTGC 420
QY 409 AGCGAGTTCTCGGCGCTTGGGTGAGAGCTATTCGGCTATGATGCGGACCAACAGA 468
Dd |||
421 AGCGAGTTCTCGGCGCTTGGGTGAGAGCTATTCGGCTATGATGCGGACCAACAGA 480
QY 469 CAATCGGCTGCTGATGCGCGCTTCCGGCTGTCAGCGAGGGCGCGGTTCTTT 528
Dd |||
481 CAATCGGCTGCTGATGCGCGCTTCCGGCTGTCAGCGAGGGCGCGGTTCTTT 540
QY 529 TTGTCGAAGCAACCTGTCGGTCCCTGAATCAACTGCAAGCAGGAGCGCGGCTAT 588
Dd |||
541 TTGTCGAAGCAACCTGTCGGTCCCTGAATCAACTGCAAGCAGGAGCGCGGCTAT 600
QY 589 CTTGGCTGCGAGCAGCGCGCTTCTTGGGAGCTGTCGTCAGCTGTCATGAGCGG 648
Dd |||
601 CTTGGCTGCGAGCAGCGCGCTTCTTGGGAGCTGTCGTCAGCTGTCATGAGCGG 660
QY 649 GAAGGACTGCTGCTATTGGGCGAGTCCCGGGCAGGATCTCTGTCATCTCACTTG 708
Dd |||
661 GAAGGACTGCTGCTATTGGGCGAGTCCCGGGCAGGATCTCTGTCATCTCACTTG 720
QY 709 CTCCTGCGAGAAAGTATCCATCATGCTGATGCAATGGCGGCTGCAATCGTGTATC 768
Dd |||
721 CTCCTGCGAGAAAGTATCCATCATGCTGATGCAATGGCGGCTGCAATCGTGTATC 780
QY 769 CGGCTACCTGCCCATTCGACCAACCAAGCGAAACATCGCATCAGCGAGCACTATCCGA 828
Dd |||
781 CGGCTACCTGCCCATTCGACCAACCAAGCGAAACATCGCATCAGCGAGCACTATCCGA 840

```

```

QY 829 TGGAGAGCCGGTCTTGTGATCAGGATGATCTGAGCAAGAGCATCAGGGGCTCGGCCAG 888
Dd |||
841 TGGAGAGCCGGTCTTGTGATCAGGATGATCTGAGCAAGAGCATCAGGGGCTCGGCCAG 900
QY 889 CCGAACTGTTCCCGAGGCTCAAGGGCGCATCCCGAGCGGAGATCTCTGTCGTGACCC 948
Dd |||
901 CCGAACTGTTCCCGAGGCTCAAGGGCGCATCCCGAGCGGAGGATCTCTGTCGTGACCC 960
QY 949 ATGGCGATGCTGCTTGGCGGAATATATGCTGGAAAAATGGCGCTTTCTGATTCATCG 1008
Dd |||
961 ATGGCGATGCTGCTTGGCGGAATATATGCTGGAAAAATGGCGCTTTCTGATTCATCG 1020
QY 1009 ACTGTGGCGGCTGGGTGCGGACCGCTATCAGACATAGCTTGGCTACCCGTGATA 1068
Dd |||
1021 ACTGTGGCGGCTGGGTGCGGACCGCTATCAGACATAGCTTGGCTACCCGTGATA 1080
QY 1069 TTGCTGAAGAGCTTGGCGGCAATGGGCTTCTCTGCTTTTACGATATCGCG 1128
Dd |||
1081 TTGCTGAAGAGCTTGGCGGCAATGGGCTTCTCTGCTTTTACGATATCGCG 1140
QY 1129 CTCGCGATTCGAGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTGAGTTTAA 1188
Dd |||
1141 CTCGCGATTCGAGCGCATCGCTTCTATCGCTTCTTACGAGTCTTCTGAGTTTAA 1200
QY 1189 CAGACCAACAAAGGTTTCCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCCCCCC 1248
Dd |||
1201 CAGACCAACAAAGGTTTCCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCCCCCC 1260
QY 1249 AACGTTACTGGCGAAGCGCTTGAATAAGGCGGTGCTGCTTGTCTATATGTTT 1308
Dd |||
1261 AACGTTACTGGCGAAGCGCTTGAATAAGGCGGTGCTGCTTGTCTATATGTTT 1320
QY 1309 TCACCATATTCGCGTCTTTGGCAATGTAGGCGCGGAAACCTTGGCCCTGCTCTCTG 1368
Dd |||
1321 TCACCATATTCGCGTCTTTGGCAATGTAGGCGCGGAAACCTTGGCCCTGCTCTCTG 1380
QY 1369 AGGAGCTTCTTGGGCTTCTTCCCTCTCGCAAGGAAATCAAGGCTCTGTTGAATGC 1428
Dd |||
1381 AGGAGCTTCTTGGGCTTCTTCCCTCTCGCAAGGAAATCAAGGCTCTGTTGAATGC 1440
QY 1429 GTGAAGGAGCAGTCTCTGGAAGCTTCTTGAAGCAAAACAGCTGTAGCGACCTT 1488
Dd |||
1441 GTGAAGGAGCAGTCTCTGGAAGCTTCTTGAAGCAAAACAGCTGTAGCGACCTT 1500
QY 1489 TGCAGGAGCGGAAACCCCGCCTCGGACAGGCTCTCGGCGCAAAAGCAGCTGTA 1548
Dd |||
1501 TGCAGGAGCGGAAACCCCGCCTCGGACAGGCTCTCGGCGCAAAAGCAGCTGTA 1560
QY 1549 TAAAGATACCTGCAAGGCGGCAAAACCCAGTGCACGTTGTGAGTTGGATAGTTG 1608
Dd |||
1561 TAAAGATACCTGCAAGGCGGCAAAACCCAGTGCACGTTGTGAGTTGGATAGTTG 1620
QY 1609 GAAAGATCAAAATGCTCTCTCAAGGTAATTCACAAAGGCTGAAGGATGCCAGAA 1668
Dd |||
1621 GAAAGATCAAAATGCTCTCTCAAGGTAATTCACAAAGGCTGAAGGATGCCAGAA 1680
QY 1669 GTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGACATGCTTACATGCTTTAG 1728
Dd |||
1681 GTACCCCATTTGATGGGATCTGATCTGGGCTCGGTGACATGCTTACATGCTTTAG 1740
QY 1729 TCGAGTTTAAAAAAGCTCTAGGCGCCCGCAACACAGGGGACGTGCTTTCTTGAATA 1788
Dd |||
1741 TCGAGTTTAAAAAAGCTCTAGGCGCCCGCAACACAGGGGACGTGCTTTCTTGAATA 1800
QY 1789 CACGATATACCATGGCGCTTATTCAGGCTTCTCCCAACAGACGCGAGGCTCTTGGC 1848
Dd |||
1801 CACGATATACCATGGCGCTTATTCAGGCTTCTCCCAACAGACGCGAGGCTCTTGGC 1860
QY 1849 TGCATCATCTAGCTCAGAGCGGAGACAGAAACAGGTCGAGGGGAGGTCCAAAGTG 1908
Dd |||
1861 TGCATCATCTAGCTCAGAGCGGAGACAGAAACAGGTCGAGGGGAGGTCCAAAGTG 1920

```

QY 1909 GTCTCCACCGCAACACAATCTTTCTGCGCAGCTCGCTCAATGGCGTGTGTGACTGTC 1968  
Db 1921 GTCTCCACCGCAACACAATCTTTCTGCGCAGCTCGCTCAATGGCGTGTGTGACTGTC 1980  
QY 1969 TATCATGGTGC CGGCTCAAAGACCTTTGCGGCGCCAAAGGCGCCAAATCACCACCAATGTAC 2028  
Db 1981 TATCATGGTGC CGGCTCAAAGACCTTTGCGGCGCCAAAGGCGCCAAATCACCACCAATGTAC 2040  
QY 2029 ACCAATGTGAC CAGGAGCTCTGTCGGCTGGGAAGGCGCCCGGGCGCGGTTCCTTGACA 2088  
Db 2041 ACCAATGTGAC CAGGAGCTCTGTCGGCTGGGAAGGCGCCCGGGCGCGGTTCCTTGACA 2100  
QY 2089 CCATGCACCTGCGGAGCTCGGACCTTTACTTTGTGTCACGAGGCATGCCGATGCTATCCG 2148  
Db 2101 CCATGCACCTGCGGAGCTCGGACCTTTACTTTGTGTCACGAGGCATGCCGATGCTATCCG 2160  
QY 2149 GTGCGCGGCGGCGGCGACAGCAGGGGAGCCCTACTCTCCCCAGGCGCCGCTCTCCTACTTG 2208  
Db 2161 GTGCGCGGCGGCGGCGGACAGCAGGGGAGCCCTACTCTCCCCAGGCGCCGCTCTCCTACTTG 2220  
QY 2209 AAGGCTCTTTCGGGCGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGCATCTTTCGG 2268  
Db 2221 AAGGCTCTTTCGGGCGTCCACTGCTCTGCCCTCGGGGATGCTGTGGGCATCTTTCGG 2280  
QY 2269 GCTGCCGTGTGCAACCCAGGSGTTTGGAAAGCGGTGGACTTTGTACCGCTGAGTCTATG 2328  
Db 2281 GCTGCCGTGTGCAACCCAGGSGTTTGGAAAGCGGTGGACTTTGTACCGCTGAGTCTATG 2340  
QY 2329 GAAACCACTATGCGGTCCCGGCTTTCACGAGCAACTCGTCCCTCGGCGGCTACCGCAG 2388  
Db 2341 GAAACCACTATGCGGTCCCGGCTTTCACGAGCAACTCGTCCCTCGGCGGCTACCGCAG 2400  
QY 2389 ACATTCAGGTGGCCATCTACAGCCCTACTGCTAGCGGCAAGAGACTAAGGTGCG 2448  
Db 2401 ACATTCAGGTGGCCATCTACAGCCCTACTGCTAGCGGCAAGAGACTAAGGTGCG 2460  
QY 2449 GCTGCTATGCAAGGCTAAGTGTCTGCTGAAACCGTCCGTCGCGCCGAC 2508  
Db 2461 GCTGCTATGCAAGGCTAAGTGTCTGCTGAAACCGTCCGTCGCGCCGAC 2520  
QY 2509 CTAGGTTTCGGGCGTATATGCTAAGGCACATGGTATCGACCCCTAAACAGAACCGG 2568  
Db 2521 CTAGGTTTCGGGCGTATATGCTAAGGCACATGGTATCGACCCCTAAACAGAACCGG 2580  
QY 2569 GTAAGGACATAC CAGGCTGCCCCCATCAGTACTCCACCTATGAGGCAAGTTCTTGCC 2628  
Db 2581 GTAAGGACATAC CAGGCTGCCCCCATCAGTACTCCACCTATGAGGCAAGTTCTTGCC 2640  
QY 2629 GACGTTGTTGCTCTGCGGCGCTATGACATCATATATATGTGATGAGTGCCACTCACT 2688  
Db 2641 GACGTTGTTGCTCTGCGGCGCTATGACATCATATATATGTGATGAGTGCCACTCACT 2700  
QY 2689 GACTCGACCACTATCTTGGGCGATCGGCACAGTCTCGACCAAGCGGAGAGCGGTGGAGCG 2748  
Db 2701 GACTCGACCACTATCTTGGGCGATCGGCACAGTCTCGACCAAGCGGAGAGCGGTGGAGCG 2760  
QY 2749 CGACTCGTCTGCTCGCCACCGTACCGCTCGGGATCGGTACCGTGCCACATPCAAAC 2808  
Db 2761 CGACTCGTCTGCTCGCCACCGTACCGCTCGGGATCGGTACCGTGCCACATPCAAAC 2820  
QY 2809 ATCGAGAGTGGTCTGTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCC 2868  
Db 2821 ATCGAGAGTGGTCTGTGTCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCC 2880  
QY 2869 ATCGAGACCACTAAGGGGGGAGGCACTCATTTTCTGCCATTCCTAAGAGAAATGTGAT 2928  
Db 2881 ATCGAGACCACTAAGGGGGGAGGCACTCATTTTCTGCCATTCCTAAGAGAAATGTGAT 2940  
QY 2929 GAGCTCGCGGAAGCTGTGCGGCTCGGACTCAATGCTGATGATATATACCGGGCCTT 2988  
Db 2941 GAGCTCGCGGAAGCTGTGCGGCTCGGACTCAATGCTGATGATATATACCGGGCCTT 3000  
QY 2989 GATGTATCCGTCATACCACTAGCGGAGAGCTCATTTGCTAGCAACGGAGCGTCTAATG 3048

Db 3001 GATGTATCCGTCATACCACTAGCGGAGAGCTCATTTGTCGTAACAACGAGCGCTTAATG 3060  
QY 3049 ACGGCTTTTACCGCGGATTTTCGACTCAGTGAATCACTGCAATATGTCACCCAGACA 3108  
Db 3061 ACGGCTTTTACCGCGGATTTTCGACTCAGTGAATCACTGCAATATGTCACCCAGACA 3120  
QY 3109 GTGCACTTACGCTTGGACCGGACCTTCACTTGAAGCGAGACCGTGCCACAGACGCG 3168  
Db 3121 GTGCACTTACGCTTGGACCGGACCTTCACTTGAAGCGAGACCGTGCCACAGACGCG 3180  
QY 3169 GTGTCACGCTCGACCGCGGAGGAGTGTAGGGGAGGATGGGCAATTTACAGTTT 3228  
Db 3181 GTGTCACGCTCGACCGCGGAGGAGTGTAGGGGAGGATGGGCAATTTACAGTTT 3240  
QY 3229 GTGCACTCCAGGAGAACCGGCGCTCGGGCATGTTCAATTCCTCGGTTCTGTGAGTGTAT 3288  
Db 3241 GTGCACTCCAGGAGAACCGGCGCTCGGGCATGTTCAATTCCTCGGTTCTGTGAGTGTAT 3300  
QY 3289 GACCGGCTGTGCTTGGTAGAGCTCAACGCGCGCGAGACCTCAGTTAGGTTGCGGCT 3348  
Db 3301 GACCGGCTGTGCTTGGTAGAGCTCAACGCGCGCGAGACCTCAGTTAGGTTGCGGCT 3360  
QY 3349 TACCTAACAACAC CAGGCTTCCCGCTCTGCCAGGACCATCTGGAGTTCTTGGGAGGCGTC 3408  
Db 3361 TACCTAACAACAC CAGGCTTCCCGCTCTGCCAGGACCATCTGGAGTTCTTGGGAGGCGTC 3420  
QY 3409 TTTACAGGCTTCAACCATAGAGCGCCATTTCTTGTCCCAGACTAAGCAGGAGGAGAC 3468  
Db 3421 TTTACAGGCTTCAACCATAGAGCGCCATTTCTTGTCCCAGACTAAGCAGGAGGAGAC 3480  
QY 3469 AACTTCCCTTACCTGTAGCATACAGGCTACGGTGTGGCGAGGCTCAGGCTTCCACT 3528  
Db 3481 AACTTCCCTTACCTGTAGCATACAGGCTACGGTGTGGCGAGGCTCAGGCTTCCACT 3540  
QY 3529 CCATCGTGGGACCAATATGTAAGTGTCTCATAGGCTAAGCTACGCTGACGCGGCA 3588  
Db 3541 CCATCGTGGGACCAATATGTAAGTGTCTCATAGGCTAAGCTACGCTGACGCGGCA 3600  
QY 3589 ACGCCCTGTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACCCCAT 3648  
Db 3601 ACGCCCTGTGTATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACCCCAT 3660  
QY 3649 ACCAATACATATGTCATGTCATGTCGCTGAGCTGAGCTGAGGCTGTCAGAGCACTGGGT 3708  
Db 3661 ACCAATACATATGTCATGTCATGTCGCTGAGCTGAGCTGAGGCTGTCAGAGCACTGGGT 3720  
QY 3709 CTGTTAGCGGAGTCTTAGCAGCTCTGGCGCGTATTTGCTGACACAGGAGCGGTGTC 3768  
Db 3721 CTGTTAGCGGAGTCTTAGCAGCTCTGGCGCGTATTTGCTGACACAGGAGCGGTGTC 3780  
QY 3769 ATTGTGGGAGGATCATTTTGTCCGAAAGCCGCGCATCATTTCCGACAGGAGTCTT 3828  
Db 3781 ATTGTGGGAGGATCATTTTGTCCGAAAGCCGCGCATCATTTCCGACAGGAGTCTT 3840  
QY 3829 TACCGGAGTTCATGAGATGGAAGTGGCTTCACTCCCTTACATGCAACAGGGA 3888  
Db 3841 TACCGGAGTTCATGAGATGGAAGTGGCTTCACTCCCTTACATGCAACAGGGA 3900  
QY 3889 ATGCACTGCCGAACAATTTCAACAGAGGCAATCGGTTGCTGCAACAGCCCAAG 3948  
Db 3901 ATGCACTGCCGAACAATTTCAACAGAGGCAATCGGTTGCTGCAACAGCCCAAG 3960  
QY 3949 CAAGCGAGGCTCTGCTCCCGTGGTGAATCCAAAGTGGCGGACCCCTCGAAGCTTCTG 4008  
Db 3961 CAAGCGAGGCTCTGCTCCCGTGGTGAATCCAAAGTGGCGGACCCCTCGAAGCTTCTG 4020  
QY 4009 GCGAAGCATATGGAATTTTCATCAGCGGATACAATTTTACAGGCTTGTCCACTCTG 4068  
Db 4021 GCGAAGCATATGGAATTTTCATCAGCGGATACAATTTTACAGGCTTGTCCACTCTG 4080  
QY 4069 CTTGGCAACCCCGGATAGCATCTGATGGGATTCACAGCTTCTATCAGGCGGCTC 4128

Db 4081 CTTGGCAACCCCGGATAGCATCATCTGATGGCATTACACAGCCTCTATCACACGCGCGTC 4140  
Qy 4129 ACCACCAACATACCTCTCTGTTTAACTCTTGGGGGATGGTGGCCGCCCAACTTGCT 4188  
Db 4141 ACCACCAACATACCTCTCTGTTTAACTCTTGGGGGATGGTGGCCGCCCAACTTGCT 4200  
Qy 4189 CTTCCAGCGCTGCTCTCTCTCTGTTAGCGCGCGCATCGCTGGAGCGGCTGTTGSCAGC 4248  
Db 4201 CTTCCAGCGCTGCTCTCTCTCTGTTAGCGCGCGCATCGCTGGAGCGGCTGTTGSCAGC 4260  
Qy 4249 ATAGGCTTTGGGAAGTGTCTTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGC 4308  
Db 4261 ATAGGCTTTGGGAAGTGTCTTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGC 4320  
Qy 4309 GCGCTCGTGGCTTTAAGTCTATGAGCGCGGAGATGCCCTCCACCGAGACCTGTTTAA 4368  
Db 4321 GCGCTCGTGGCTTTAAGTCTATGAGCGCGGAGATGCCCTCCACCGAGACCTGTTTAA 4380  
Qy 4369 CTACTCCCTGCTATCCTCTCCCTGGCGCCCTAGTCTGGGGTCTGTGGCAGCGATA 4428  
Db 4381 CTACTCCCTGCTATCCTCTCCCTGGCGCCCTAGTCTGGGGTCTGTGGCAGCGATA 4440  
Qy 4429 CTGCGTGGCAGCTGCGGCCAGGGGAGGGGCTGTGCATGTGATGAACCGGCTGATAGCG 4488  
Db 4441 CTGCGTGGCAGCTGCGGCCAGGGGAGGGGCTGTGCATGTGATGAACCGGCTGATAGCG 4500  
Qy 4489 TTGCGTTCCGGGGTAAACAGCTCTCCCGCAGCAGTATGTGCTGAGAGCGAGCTGCA 4548  
Db 4501 TTGCGTTCCGGGGTAAACAGCTCTCCCGCAGCAGTATGTGCTGAGAGCGAGCTGCA 4560  
Qy 4549 GCAGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGTGTCTGAAGAGCTTCAC 4608  
Db 4561 GCAGTGTCACTCAGATCCTCTCTAGTCTTACCATCACTCAGTGTCTGAAGAGCTTCAC 4620  
Qy 4609 CAGTGGATCAACGAGAGCTCTCCAGCCATGCTCCGGTCTGTTGCTAAGAGATGTTGG 4668  
Db 4621 CAGTGGATCAACGAGAGCTCTCCAGCCATGCTCCGGTCTGTTGCTAAGAGATGTTGG 4680  
Qy 4669 GATTGATATGACGCTGTGATGATTTCAAGACCTGGCTCCAGTCCAGCTCTCCGCG 4728  
Db 4681 GATTGATATGACGCTGTGATGATTTCAAGACCTGGCTCCAGTCCAGCTCTCCGCG 4740  
Qy 4729 CGATTCCGGGAGTCCCTCTCTCTCATGTCAACGTGGGTCAAGAGGAGTCTGGCGGGG 4788  
Db 4741 CGATTCCGGGAGTCCCTCTCTCTCATGTCAACGTGGGTCAAGAGGAGTCTGGCGGGG 4800  
Qy 4789 GACGGCATCATGCAAAACCACTGCCCCATGTGGAGCACAGATCACCGGACATGTGAAAAA 4848  
Db 4801 GACGGCATCATGCAAAACCACTGCCCCATGTGGAGCACAGATCACCGGACATGTGAAAAA 4860  
Qy 4849 GGTTCATGAGGATCGTGGGCTAGACCTGTAGTAAACGTTGGCATGGACATTCGCC 4908  
Db 4861 GGTTCATGAGGATCGTGGGCTAGGACCTGTAGTAAACGTTGGCATGGACATTCGCC 4920  
Qy 4909 ATTAAACGCTFACACCAACCGGCGCTGACGCCCTCCCGCGCCAAATATTCTAGGGG 4968  
Db 4921 ATTAAACGCTFACACCAACCGGCGCTGACGCCCTCCCGCGCCAAATATTCTAGGGG 4980  
Qy 4969 CTGTGCGGGTGTCTGTAGAGAGTACGTGGAGTTACGCGGTGGGGATTTCCACTAC 5028  
Db 4981 CTGTGCGGGTGTCTGTAGAGAGTACGTGGAGTTACGCGGTGGGGATTTCCACTAC 5040  
Qy 5029 GTACCGGGATGACCACTGACACGTAAGTGGCCGCTGTCAGTTCCGGCCCGGAATTC 5088  
Db 5041 GTACCGGGATGACCACTGACACGTAAGTGGCCGCTGTCAGTTCCGGCCCGGAATTC 5100  
Qy 5089 TTCAACAGAGTGGATGGGTCGCTGCAACAGTACGCTCCAGCGTGCAAAACCCCTCCTA 5148  
Db 5101 TTCAACAGAGTGGATGGGTCGCTGCAACAGTACGCTCCAGCGTGCAAAACCCCTCCTA 5160  
Qy 5149 CGGAGAGGTCACTTCCTGCTGGGCTCAATCAATACCTGTTGGGTCAAGTCCCA 5208  
Db 5161 CGGAGAGGTCACTTCCTGCTGGGCTCAATCAATACCTGTTGGGTCAAGTCCCA 5220

Qy 5209 TCGGAGCCGACCGGACGTAGCAGTGTCTCATCTCATGCTCACGACCCCTCCCACTT 5268  
Db 5221 TCGGGCCCCGAACCGGACGTAGCAGTGTCTCATCTCATGCTCACGACCCCTCCCACTT 5280  
Qy 5269 ACGGGGAGACCGGCTAAGCGTAGGCTGGCCAGGGATCTCCCCCTCTTTGGCCAGCTCA 5328  
Db 5281 ACGGGGAGACCGGCTAAGCGTAGGCTGGCCAGGGATCTCCCCCTCTTTGGCCAGCTCA 5340  
Qy 5329 TCAGCTAGCAGCTGTCTGGGCTTCTTTGAAGGCAACATGACACTACCCGTATGACTCC 5388  
Db 5341 TCAGCTAGCAGCTGTCTGGGCTTCTTTGAAGGCAACATGACACTACCCGTATGACTCC 5400  
Qy 5389 CGGAGCGTGAATCATCGAGCCCAACCTCTGTGGCGCAGAGATGGCGGGAACATC 5448  
Db 5401 CGGAGCGTGAATCATCGAGCCCAACCTCTGTGGCGCAGAGATGGCGGGAACATC 5460  
Qy 5449 ACCCGGTGGAGTCAAGAAATAGGTAGTAAATTTTGGACTTTTCGAGCGCTTCCAAAGCG 5508  
Db 5461 ACCCGGTGGAGTCAAGAAATAGGTAGTAAATTTTGGACTTTTCGAGCGCTTCCAAAGCG 5520  
Qy 5509 GAGGAGGATGAGGGAAGTATCCGTTCCGGCGGAGATCCTCGGAGGTCCAGGAATTC 5568  
Db 5521 GAGGAGGATGAGGGAAGTATCCGTTCCGGCGGAGATCCTCGGAGGTCCAGGAATTC 5580  
Qy 5569 CCTCGAGCGATGCCATATGGGCACGCCCGGATTAACAACCTCCACTGTTAGAGTCTCTGG 5628  
Db 5581 CCTCGAGCGATGCCATATGGGCACGCCCGGATTAACAACCTCCACTGTTAGAGTCTCTGG 5640  
Qy 5629 AAGGACCCGAGTACGTCTCCCTCAGTGTACACGGGTTCATTTGCGCGCTGTCGAAGGCC 5688  
Db 5641 AAGGACCCGAGTACGTCTCCCTCAGTGTACACGGGTTCATTTGCGCGCTGTCGAAGGCC 5700  
Qy 5689 CTTCCGATACCACTCCAGGAGGAAGAGACGGTTGTCTGTCTGTCTCAGAACTACCCGTCT 5748  
Db 5701 CTTCCGATACCACTCCAGGAGGAGAGACGGTTGTCTGTCTGTCTCAGAACTACCCGTCT 5760  
Qy 5749 TCTGCTTCGGCGAGTCTGCCCAAGACCTTCGGCAGCTCCGAAATCGTCGGCGCTCGAC 5808  
Db 5761 TCTGCTTCGGCGAGTCTGCCCAAGACCTTCGGCAGCTCCGAAATCGTCGGCGCTCGAC 5820  
Qy 5809 AGCGGACCGGCAACCGCTCTCTGACACGCGCTTCGACGACGGCGACGCGGATCCGAC 5868  
Db 5821 AGCGGACCGGCAACCGCTCTCTGACACGCGCTTCGACGACGCGGATCCGAC 5880  
Qy 5869 GTTGAGTCTGTCTCTCATGCCCCCTTTGAGGGGAGCGCGGGATCCCGATCTCAGC 5928  
Db 5881 GTTGAGTCTGTCTCTCATGCCCCCTTTGAGGGGAGCGCGGGATCCCGATCTCAGC 5940  
Qy 5929 GACGGTCTTGGTCTACCGTAAGCAGGAGGCTTAGTGAGGACGTCTGTCTGTCTCGATG 5988  
Db 5941 GACGGTCTTGGTCTACCGTAAGCAGGAGGCTTAGTGAGGACGTCTGTCTGTCTCGATG 6000  
Qy 5989 TCTTACATGACAGCGCGCTGTATCAGCCATGCGCTGCGGAGGAAACCAAGCTGCC 6048  
Db 6001 TCTTACATGACAGCGCGCTGTATCAGCCATGCGCTGCGGAGGAAACCAAGCTGCC 6060  
Qy 6049 ATCAATGCACTGAGCAACTCTTTGCTCCGTCCACCACTTGTGTCTATGCTTACCAATCT 6108  
Db 6061 GTCAATGCACTGAGCAACTCTTTGCTCCGTCCACCACTTGTGTCTATGCTTACCAATCT 6120  
Qy 6109 CGCAGCGCAAGCTCGCGCAGAGAGGTCACTTTTGAAGACTGACAGTCTCTGACGAC 6168  
Db 6121 CGCAGCGCAAGCTCGCGCAGAGAGGTCACTTTTGAAGACTGACAGTCTCTGACGAC 6180  
Qy 6169 CACTACCGGAGCGTCTCAAGGAGATGAAGCGGAGCGCTTCCAGTTAAGGTAACTT 6228  
Db 6181 CACTACCGGAGCGTCTCAAGGAGATGAAGCGGAGCGCTTCCAGTTAAGGTAACTT 6240  
Qy 6229 CTAATCGTGGAGGAGCTGTAGCTGACGCCCCACATTCGCGCAGATCTAATTTGGC 6288  
Db 6241 CTAATCGTGGAGGAGCTGTAGCTGACGCCCCACATTCGCGCAGATCTAATTTGGC 6300



6289 TATGGGCAAGACGCTCGGAACCTATCCAGAAAGCCGTTAAACACATCCGCTCGGTG 6348  
6301 TATGGGCAAGACGCTCGGAACCTATCCAGAAAGCCGTTAAACACATCCGCTCGGTG 6360  
6349 TGAAGGACCTTGTCTGAAGACACTGAGACACCAATTAACACACCAATCATGGCAAAAAT 6408  
6361 TGAAGGACCTTGTCTGAAGACACTGAGACACCAATTAACACACCAATCATGGCAAAAAT 6420  
6409 GAGTTTTCGCTCCAAACAGAGAAAGGGGGCGCAAGCAGCTCGCTTATCGTATTC 6468  
6421 GAGTTTTCGCTCCAAACAGAGAAAGGGGGCGCAAGCAGCTCGCTTATCGTATTC 6480  
6469 CAGATTGGGGGTCGTGTGCGGAATAAGCCCTTTACGATGCGTCTCCACCTC 6528  
6481 CAGATTGGGGGTCGTGTGCGGAATAAGCCCTTTACGATGCGTCTCCACCTC 6540  
6529 CCTCAGCCGTCATGGCTCTTCATACGGAATCCAAATCTCTCTGACAGCGGTCGAG 6588  
6541 CCTCAGCCGTCATGGCTCTTCATACGGAATCCAAATCTCTCTGACAGCGGTCGAG 6600  
6589 TTCCTGTAATGCTGGAAGCGAAGAAATGCCCTTATGGCTTCGCATATGACACCCGC 6648  
6601 TTCCTGTAATGCTGGAAGCGAAGAAATGCCCTTATGGCTTCGCATATGACACCCGC 6660  
6649 TGTCTTACCTCAAGGTCACCTGAGATGACATCGCTTTCGAGGATCAATCTACCAATGT 6708  
6661 TGTCTTACCTCAAGGTCACCTGAGATGACATCGCTTTCGAGGATCAATCTACCAATGT 6720  
6709 TGTGACTTGGCCCCCGAAGCAGACAGGCCATTAAGTTCGCTCAGAGCGGCTTTATATC 6768  
6721 TGTGACTTGGCCCCCGAAGCAGACAGGCCATTAAGTTCGCTCAGAGCGGCTTTATATC 6780  
6769 GGGGGCCCCCTGACTAATTTAAAGGCGAAGCTGCGGCTATCGCGGTCGCGCGGAGC 6828  
6781 GGGGGCCCCCTGACTAATTTAAAGGCGAAGCTGCGGCTATCGCGGTCGCGCGGAGC 6840  
6829 GGTGTACTGACGACACGCTGGGTATATACCTCACAATGTTTAAAGCGGCTGCGGCC 6888  
6841 GGTGTACTGACGACACGCTGGGTATATACCTCACAATGTTTAAAGCGGCTGCGGCC 6900  
6889 TGTGAGCTCGAAGCTTCCAGGACTGACGATGCTGCTATGCGGAGACGACCTTGTGCTT 6948  
6901 TGTGAGCTCGAAGCTTCCAGGACTGACGATGCTGCTATGCGGAGACGACCTTGTGCTT 6960  
6949 ATCTGAAAGCGCGGGAACCAAGAGGACGAGCGCTACGGGCTTTCACGAGGCT 7008  
6961 ATCTGAAAGCGCGGGAACCAAGAGGACGAGCGCTACGGGCTTTCACGAGGCT 7020  
7009 ATGACTAGATACCTGCCCCCTGGGGACCGCCCAACCAAGATACGACTTGGAGTTG 7068  
7021 ATGACTAGATACCTGCCCCCTGGGGACCGCCCAACCAAGATACGACTTGGAGTTG 7080  
7069 ATAACATCATGCTCCTCCAATGTCAGTCGCGCAGCATGATCTGGCAAAAAGGCTGAC 7128  
7081 ATAACATCATGCTCCTCCAATGTCAGTCGCGCAGCATGATCTGGCAAAAAGGCTGAC 7140  
7129 TATCTACCGCTGACCCCAACACCCCTTTCGCGGGCTGCGTGGGAGACAGCTAGACAC 7188  
7141 TATCTACCGCTGACCCCAACACCCCTTTCGCGGGCTGCGTGGGAGACAGCTAGACAC 7200  
7189 ACTCCAGTCAATCTCGCTAGGCAACATCATGATGATGGCCCACTTGTGGGCAAGG 7248  
7201 ACTCCAGTCAATCTCGCTAGGCAACATCATGATGATGGCCCACTTGTGGGCAAGG 7260  
7249 ATGATCTGATGACTCATTTCTTCTCCATCTTCTAGCTCAGGAACCACTTGA AAAAGCC 7308  
7261 ATGATCTGATGACTCATTTCTTCTCCATCTTCTAGCTCAGGAACCACTTGA AAAAGCC 7320  
7309 CTGATTGTGAGATCTACGGGGCTGTACTTCCATTGAGCCACTTACCTCAGATC 7368  
7321 CTGATTGTGAGATCTACGGGGCTGTACTTCCATTGAGCCACTTACCTCAGATC 7380  
7369 ATTCAACGACTCCATGGCTTAGCGCAATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7428

7381 ATTCAACGACTCCATGGCCTTAGCGCAATTTTCACTCCATAGTTACTCTCCAGGTGAGATC 7440  
7429 AATAGGCTGGCTTTCATGCTCCTCAGAAACCTTGGGGTACCGCCCTTGGAGTCTCGAGACAT 7488  
7441 AATAGGCTGGCTTTCATGCTCCTCAGAAACCTTGGGGTACCGCCCTTGGAGTCTCGAGACAT 7500  
7489 CGGGCAGAAAGTTCGCGCTAGGCTACTGTCCAGGGGGGAGGGTGCACCTTGTGGC 7548  
7501 CGGGCAGAAAGTTCGCGCTAGGCTACTGTCCAGGGGGGAGGGTGCACCTTGTGGC 7560  
7549 AAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7608  
7561 AAGTACTCTTCAACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7620  
7609 TCCAGTTGGAATTTATCCAGCTGTTTCTGTTGCTGTTTACAGCGGGGAGACATATATCAC 7668  
7621 TCCAGTTGGAATTTATCCAGCTGTTTCTGTTGCTGTTTACAGCGGGGAGACATATATCAC 7680  
7669 AGCTGTCTGTCGTCGACCCCGCTGTTTCACTGTTGCTGTTTACAGCGGGGAGACATATATCAC 7728  
7681 AGCTGTCTGTCGTCGACCCCGCTGTTTCACTGTTGCTGTTTACAGCGGGGAGACATATATCAC 7740  
7729 GTAGGCACTATCTACTTCCCAACCGATGAACGGGACCTTAAACACTCCAGGCCAATAGG 7788  
7741 GTAGGCACTATCTACTTCCCAACCGATGAACGGGACCTTAAACACTCCAGGCCAATAGG 7800  
7789 CCATCTGTTTTTCCCTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7848  
7801 CCATCTGTTTTTCCCTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860  
7849 TTTTCTCTTTTTTTTTTCTT 7908  
7861 TTTTCTCTTTTTTTTTTCTT 7920  
7909 CCTAGTCAGCGTAGCTGTGAAGGTCGTAAGGTCGTAAGGTCGTAAGGTCGTAAGGTCGTAAGG 7968  
7921 CCTAGTCAGCGTAGCTGTGAAGGTCGTAAGGTCGTAAGGTCGTAAGGTCGTAAGGTCGTAAGG 7980  
7969 TGGCCTCTCTGCAGATCAAGT 7989  
7981 TGGCCTCTCTGCAGATCAAGT 8001

RESULT 5  
US-09-539-601-28  
; Sequence 28, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; CURRENT FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 28  
; LENGTH: 8001  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct 1389/NS3-3'/19  
; FEATURE: CDS  
; NAME/KEY: CDS  
; LOCATION: (342)..(1193)  
; OTHER INFORMATION: hepatitis C virus core - neomycin  
; FEATURE: phosphotransferase fusion protein  
; NAME/KEY: RBS

```

; LOCATION: (1202)..(1812)
; OTHER INFORMATION: internal ribosome entry site from
; OTHER INFORMATION: encephalomyocarditis virus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1813)..(7770)
; OTHER INFORMATION: hepatitis C virus nonstructural proteins N83 - 5B
; OTHER INFORMATION: of cell culture adapted clone no. 19
; FEATURE:
; NAME/KEY: 3'UTR
; LOCATION: (7771)..(8001)
; US-09-539-601-28

Query Match          99.4%; Score 7939.8; DB 4; Length 8001;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 7972; Conservative 0; Mismatches 17; Indels 12; Gaps 1;

QY 1 GCCAGCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
DB 1 GCCAGCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60

QY 61 TCTTCACGAGAAAGGCTTAGCCATGGGCTTAGTATGATGATGCTGCGAGCCTCCAGGAC 120
DB 61 TCTTCACGAGAAAGGCTTAGCCATGGGCTTAGTATGATGATGCTGCGAGCCTCCAGGAC 120

QY 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAG 180

QY 181 GAGACACCGGCTCTTCTTGGATCAACCCGCTCAATGCTGAGATTTGGGGCTGCCCC 240
DB 181 GAGACACCGGCTCTTCTTGGATCAACCCGCTCAATGCTGAGATTTGGGGCTGCCCC 240

QY 241 GCGAGACTCTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGGTACTGCTGATAGG 300
DB 241 GCGAGACTCTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGGTACTGCTGATAGG 300

QY 301 GTGCTTGGAGTCCCGGAGGCTCTGAGACCGGTGACCATGAGCAGCAATCTTAAC 360
DB 301 GTGCTTGGAGTCCCGGAGGCTCTGAGACCGGTGACCATGAGCAGCAATCTTAAC 360

QY 361 CTCAAGAAAAACCCAAA-----GGCGCGCCATGATTGAACAAGATGATTGC 408
DB 361 CTCAAGAAAAACCCAAA-----GGCGCGCGCCATGATTGAACAAGATGATTGC 420

QY 409 AGCGAGTTCTCGCGCGCTTGGGTGAGAGAGCTATTGGCTATGACTGGGCACACAGA 468
DB 409 AGCGAGTTCTCGCGCGCTTGGGTGAGAGAGCTATTGGCTATGACTGGGCACACAGA 480

QY 469 CAATCGGCTGCTCTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGCCGGTTCTTT 528
DB 481 CAATCGGCTGCTCTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGCCGGTTCTTT 540

QY 529 TTGTCAAGACCACTGTCCGGTCCCTGAATGAATGAGACAGGAGCGCGGCTAT 588
DB 541 TTGTCAAGACCACTGTCCGGTCCCTGAATGAATGAGACAGGAGCGCGGCTAT 600

QY 589 CGTGCTGCGCACGAGCGGCTTCTTGGCGAGCTGTGCTCGACGTTGTCACTGAGCGG 648
DB 601 CGTGCTGCGCACGAGCGGCTTCTTGGCGAGCTGTGCTCGACGTTGTCACTGAGCGG 660

QY 649 GAAGGACTGGCTGCTATTGGGCGAAGTCCCGGGGAGGATCTCCTGTCTCATCTCACTTG 708
DB 661 GAAGGACTGGCTGCTATTGGGCGAAGTCCCGGGGAGGATCTCCTGTCTCATCTCACTTG 720

QY 709 CTCCTGCGAGAAAGTATCCATCATGCTGATGCAATGCGGGCTGATAGCTTGATC 768
DB 721 CTCCTGCGAGAAAGTATCCATCATGCTGATGCAATGCGGGCTGATAGCTTGATC 780

QY 769 CGGCTACCTGCCCATTCGACACCAAGCGAATCATCGCATCGAGCGACCTACTCGGA 828
DB 781 CGGCTACCTGCCCATTCGACACCAAGCGAATCATCGCATCGAGCGACCTACTCGGA 840

```

```

829 TCGAAGCCGGTCTTGTCTGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCCGAG 888
DB TCGAAGCCGGTCTTGTCTGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCCGAG 900

QY 889 CCGAACTGTTTCGCCAGGCTCAAGGCGCGCATCCCGACGCGGAGGATCTCGTGTGACCC 948
DB CCGAACTGTTTCGCCAGGCTCAAGGCGCGCATCCCGACGCGGAGGATCTCGTGTGACCC 960

QY 949 ATGGGATGCTGCTGTGCGGAATATCATGCTGGAATAATGCGCGCTTCTTGTGATTCATCG 1008
DB ATGGGATGCTGCTGTGCGGAATATCATGCTGGAATAATGCGCGCTTCTTGTGATTCATCG 1020

QY 1009 ACTGTGGCGGCTGGGTGCGGACCGCTATCAGGACATAGCTGTTGGCTACCGTGATA 1068
DB ACTGTGGCGGCTGGGTGCGGACCGCTATCAGGACATAGCTGTTGGCTACCGTGATA 1080

QY 1069 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTCTGCTGCTTTACGATTTAA 1128
DB TTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTCTGCTGCTTTACGATTTAA 1140

QY 1129 CTCGCGATTCGAGCGCATCGCTTCTATCGGCTTCTTACGAGTCTTCTGAGTTAA 1188
DB CTCGCGATTCGAGCGCATCGCTTCTATCGGCTTCTTACGAGTCTTCTGAGTTAA 1200

QY 1189 CAGACCAAAACGGTTCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCCCCCCCT 1248
DB CAGACCAAAACGGTTCCTCTAGCGGATCAATTCGCGCTTCTCCCTCCCCCCCCCT 1260

QY 1249 AAGCTTATGCGCGAAGCGCTTGAATAAGGCGGCTGCTTGTCTATATGTTATTT 1308
DB AAGCTTATGCGCGAAGCGCTTGAATAAGGCGGCTGCTTGTCTATATGTTATTT 1320

QY 1309 TCCACCATATGCGCTTCTTGGCAATGAGGCGCGGAAACCTGCGCTCTCTCTTG 1368
DB TCCACCATATGCGCTTCTTGGCAATGAGGCGCGGAAACCTGCGCTCTCTCTTG 1380

QY 1369 ACGAGCTTCTAGGGGCTTTCCTCTCGCAAGGAATCAAGGCTCTGTTGAATGTC 1428
DB ACGAGCTTCTAGGGGCTTTCCTCTCGCAAGGAATCAAGGCTCTGTTGAATGTC 1440

QY 1429 GTGAAGGAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACAGCTGTAGCGACCTT 1488
DB GTGAAGGAGCAGTTCCTCTGGAAGCTTCTTGAAGCAAAACAGCTGTAGCGACCTT 1500

QY 1489 TCGAGGACGCGAAGCCCGCTTGGCGACAGGCTCTGCGGCAAGGCAAGCTGTA 1548
DB TCGAGGACGCGAAGCCCGCTTGGCGACAGGCTCTGCGGCAAGGCAAGCTGTA 1560

QY 1549 TAAGATACACCTGCAAGCGGCAACCCCGAGTGCACGTTGTGAGTTGATGTTG 1608
DB TAAGATACACCTGCAAGCGGCAACCCCGAGTGCACGTTGTGAGTTGATGTTG 1620

QY 1609 GAAAGAGTCAATGCTCTCTCAAGCGTATTCACAAAGGGGCTCAAGGATGCCAGAAG 1668
DB GAAAGAGTCAATGCTCTCTCAAGCGTATTCACAAAGGGGCTCAAGGATGCCAGAAG 1680

QY 1669 GTACCCCATCTATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGCTTTAG 1728
DB GTACCCCATCTATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGCTTTAG 1740

QY 1729 TCGAGTTAAAAAAGCTCTAGGCCCCCGCAACACGCGGAGCTGTTTCTTTGAAAAA 1788
DB TCGAGTTAAAAAAGCTCTAGGCCCCCGCAACACGCGGAGCTGTTTCTTTGAAAAA 1800

QY 1789 CACGATAATACATGCGGCTTATACGGCTTACTCCCAACAGACGCGAGGCTTCTGCG 1848
DB CACGATAATACATGCGGCTTATACGGCTTACTCCCAACAGACGCGAGGCTTCTGCG 1860

QY 1849 TGCATCATCTAGCTCACAGCGCGGACAGGACAGGCTCGAGGGGAGTCCAAAGT 1908
DB TGCATCATCTAGCTCACAGCGCGGACAGGACAGGCTCGAGGGGAGTCCAAAGT 1920

QY 1909 GTCTCCCGCAACACAAATCTTTCTGCGACCTCGTCAATGGGCTGTGTTGGATGTC 1968

```

1921 GTCTCCACGCAACACATCTTTCTGGGACCTGCTGTAATGGCGTGTGTGGACGTGC 1980  
1969 TATCATGGTGCCGGCTCAAGACCCCTTGGCGGCCCAAGGGCCCAATCACCCAAAATGTAC 2028  
1981 TATCATGGTGCCGGCTCAAGACCCCTTGGCGGCCCAAGGGCCCAATCACCCAAAATGTAC 2040  
2029 ACCAATGTGGACAGGACCTCGTGGGTGCGAGCGGCCCGGGGGCGGCTTCTGTGACA 2088  
2041 ACCAATGTGGACAGGACCTCGTGGGTGCGAGCGGCCCGGGGGCGGCTTCTGTGACA 2100  
2089 CCATGACCTGCGCGGAGCTCGGACCTTTACTTGTACAGGAGCATGCCGATGTCATTTCGG 2148  
2101 CCATGACCTGCGCGGAGCTCGGACCTTTACTTGTACAGGAGCATGCCGATGTCATTTCGG 2160  
2149 GTGCGCGCGGGGGGACAGAGGGGAGCTTACTCTCCCGCAGGCCGCTCTCTACTTG 2208  
2161 GTGCGCGCGGGGGGACAGAGGGGAGCTTACTCTCTCCCGCAGGCCGCTCTCTACTTG 2220  
2209 AAGGGCTCTTCGGGGCGGTGCACCTGCTCTGCCCTCGGGGACCGTGTGGGCATCTTTCGG 2268  
2221 AAGGGCTCTTTCGGGGCGGTGCACCTGCTCTGCCCTCGGGGACCGTGTGGGCATCTTTCGG 2280  
2269 GCTGCGGTGTGCACCCGAGGGGTGCGAAGCGGTGGACTTTGTACCCCGTCGAGTCTATG 2328  
2281 GCTGCGGTGTGCACCCGAGGGGTGCGAAGCGGTGGACTTTGTACCCCGTCGAGTCTATG 2340  
2329 GAAACCACTATGCGGTCCCGGTCTTTCACGGGCAACTCGTCCCTCCGGGCCGTACCGCAG 2388  
2341 GGAACCACTATGCGGTCCCGGTCTTTCACGGGCAACTCGTCCCTCCGGGCCGTACCGCAG 2400  
2389 ACATTCAGGTGCGCCATCTACAGGCCCTACTGTGTAGCGGGAAGACACTAAGTGCCTG 2448  
2401 ACATTCAGGTGCGCCATCTACAGGCCCTACTGTGTAGCGGGAAGACACTAAGTGCCTG 2460  
2449 GCTGCGGTATGACGCCAAGGGTAAAGTGTCTGCTGAAACCGTCCGTCGCGGCCACC 2508  
2461 GCTGCGGTATGACGCCAAGGGTAAAGTGTCTGCTGAAACCGTCCGTCGCGGCCACC 2520  
2509 CTAGGTTTCGGGGCGTATATGTCTAAGGCACATGGTATCGACCTCAACATCAGAACCGGG 2568  
2521 CTAGGTTTCGGGGCGTATATGTCTAAGGCACATGGTATCGACCTCAACATCAGAACCGGG 2580  
2569 GTAAGACCACTACACCGGTGCCCCCATCAGGTACTCCACCTATGTCGAAGTTTCTTGCC 2628  
2581 GTAAGACCACTACACCGGTGCCCCCATCAGGTACTCCACCTATGTCGAAGTTTCTTGCC 2640  
2629 GACGGTGTGTCTGCGGGCGCTATGACATCATATATATGTATGATGAGTGCCTCAACT 2688  
2641 GACGGTGTGTCTGCGGGCGCTATGACATCATATATATGTATGATGAGTGCCTCAACT 2700  
2689 GACTCGACCACTATCTCTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGCGGTGGAGCG 2748  
2701 GACTCGACCACTATCTCTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGCGGTGGAGCG 2760  
2749 CGACTGCTGCTCGGCACCGCTACGCTTCGGGATCGGTACCGTGCACATCCAAAC 2808  
2761 CGACTGCTGCTGCTCGGCACCGCTACGCTTCGGGATCGGTACCGTGCACATCCAAAC 2820  
2809 ATCGAGGAGTGGTCTGTCCGACCTGGAGAAATCCCTTTTATGCGAAAGCCATCCCC 2868  
2821 ATCGAGGAGTGGTCTGTCCGACCTGGAGAAATCCCTTTTATGCGAAAGCCATCCCC 2880  
2869 ATCGAGACCACTAAGGGGGGAGGCACCTCATTTTCTGCCATTCCAAAGAGAAATGTGAT 2928  
2881 ATCGAGACCACTAAGGGGGGAGGCACCTCATTTTCTGCCATTCCAAAGAGAAATGTGAT 2940  
2929 GAGCTCGCGGAAGCTGTCTCGGCTCGGACTCAATGCTGTAGCATATATACCGGGGCTT 2988  
2941 GAGCTCGCGGAAGCTGTCTCGGCTCGGACTCAATGCTGTAGCATATATACCGGGGCTT 3000  
2989 GATGTATCCGTATACCACTAGCGGAGCGTCAATCTGCTGAGCAACGAGCGCTCTAATG 3048

3001 GATGTATCCGTATACCACTAGCGGAGAGCTCATTTGCTAGTCAACGAGCGCTCTAATG 3060  
3049 ACGGGCTTTACGGCGGATTTGCACTCAGTGATCGACTGCAATGTCATGTGTCAACGAGACA 3108  
3061 ACGGGCTTTACGGCGGATTTGCACTCAGTGATCGACTGCAATGTCATGTGTCAACGAGACA 3120  
3109 GTCGACTTCAGCTGAGCCGACCTTACCATTTGAGACGACGCGCTGCGCACAGAGCGG 3168  
3121 GTCGACTTCAGCTGAGCCGACCTTACCATTTGAGACGACGCGCTGCGCACAGAGCGG 3180  
3169 GTCTCACTCGCAGCGGGGAGGACGACTGTGTAGGGGAGGATGGGCAATTTACAGGTTT 3228  
3181 GTGTCACTCGCAGCGGGGAGGACGACTGTGTAGGGGAGGATGGGCAATTTACAGGTTT 3240  
3229 GTCACTTCAGGAGAAACGGCCCTCGGCAATTTTCGATTTCTCTCGGTCTGTGCGAGTGTAT 3288  
3241 GTCACTTCAGGAGAAACGGCCCTCGGCAATTTTCGATTTCTCTCGGTCTGTGCGAGTGTAT 3300  
3289 GACGCGGCTGTCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGGTTTGGGGCT 3348  
3301 GACGCGGCTGTCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGGTTTGGGGCT 3360  
3349 TACTTAAACACACACAGCGGTTGCCGCTGTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTC 3408  
3361 TACTTAAACACACAGCGGTTGCCGCTGTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTC 3420  
3409 TTTACAGGCTCACCCACATAGACGCGCCATTTCTTGTCCAGACTAAGAGGAGGAGAC 3468  
3421 TTTACAGGCTCACCCACATAGACGCGCCATTTCTTGTCCAGACTAAGAGGAGGAGAC 3480  
3469 AACTTCCCCTACCTGTAGCATACGAGCTACCGTGTGCGCCAGGCTCAGGCTCCACCT 3528  
3481 AACTTCCCCTACCTGTAGCATACGAGCTACCGTGTGCGCCAGGCTCAGGCTCCACCT 3540  
3529 CCATCTGGGACCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTACGCTGCACGGGCCA 3588  
3541 CCATCTGGGACCAAAATGTGAAAGTGTCTCATACGGCTAAAGCTACGCTGCACGGGCCA 3600  
3589 ACGCCCTGCTGTATAGGCTGGGAGCGGTTTCAAAAAGAGTTACTACACACACCCCATTA 3648  
3601 ACGCCCTGCTGTATAGGCTGGGAGCGGTTTCAAAAAGAGTTACTACACACACCCCATTA 3660  
3649 ACCAAATACATCATGCGATGCTGCGTGCCTGAGGTCGCTCAGGAGCTCAGGAGCCTGGTG 3708  
3661 ACCAAATACATCATGCGATGCTGCGTGCCTGAGGTCGCTCAGGAGCTCAGGAGCCTGGTG 3720  
3709 CTGCTAGGCGGAGTCTTAGCAGCTCTGGCGCGCTATTGCTCTGACAAACAGCAGCGTGGTC 3768  
3721 CTGCTAGGCGGAGTCTTAGCAGCTCTGGCGCGCTATTGCTCTGACAAACAGCAGCGTGGTC 3780  
3769 ATTGTGGCAGGATCATCTTTGTCCGAAAGCCGGCCATCATTTCCCGACAGGGAAGTCTCTT 3828  
3781 ATTGTGGCAGGATCATCTTTGTCCGAAAGCCGGCCATCATTTCCCGACAGGGAAGTCTCTT 3840  
3829 TACCGGAGTTCATGAGATGGAAGAGTGGCTCAGCTCAGCTCCCTTACATCATGCAACAGGA 3888  
3841 TACCGGAGTTCATGAGATGGAAGAGTGGCTCAGCTCAGCTCCCTTACATCATGCAACAGGA 3900  
3889 ATGAGCTCTCCCGAAACAAATTCAAACAGAGGCAATCGGGTTGTGCAAAACAGGCCACCAAG 3948  
3901 ATGAGCTCTCCCGAAACAAATTCAAACAGAGGCAATCGGGTTGTGCAAAACAGGCCACCAAG 3960  
3949 CAAGCGGAGCTGCTGCTCCCGTGGTGAATCAAAGTGGCGGACCATCGAAGCCTTCTGG 4008  
3961 CAAGCGGAGCTGCTGCTCCCGTGGTGAATCAAAGTGGCGGACCATCGAAGCCTTCTGG 4020  
4009 GCGAAGCATATGGAATTTTCATCAGGGGATCAATATTTAGCAGGCTTGTCCACTCTG 4068  
4021 GCGAAGCATATGGAATTTTCATCAGGGGATCAATATTTAGCAGGCTTGTCCACTCTG 4080  
4069 CTGTGGCAACCCCGGATAGCATCACTGATGGCAATTCACAGCTCTATCAGCAGCGCGCTC 4128  
4081 CTGTGGCAACCCCGGATAGCATCACTGATGGCAATTCACAGCTCTATCAGCAGCGCGCTC 4140

QY 4129 ACCACCAACATACCCCTCTGTTTAAATCATCTGGGGGATGGTGGCCGCCCAACTTGCT 4188  
DB 4141 ACCACCAACATACCCCTCTGTTTAAATCATCTGGGGGATGGTGGCCGCCCAACTTGCT 4200  
QY 4189 CTTCCAGCGCTCTCTCTCTTTGTTAGCGCGCGCATCGCTGGAGCGCTGTGTGSCAGC 4248  
DB 4201 CTTCCAGCGCTCTCTCTCTTTGTTAGCGCGCGCATCGCTGGAGCGCTGTGTGSCAGC 4260  
QY 4249 ATAGGCTTTGGAAAGTGTGTGGATATTTTGGCAGGTTATGAGCAGGGTGGCAGGC 4308  
DB 4261 ATAGGCTTTGGAAAGTGTGTGGATATTTTGGCAGGTTATGAGCAGGGTGGCAGGC 4320  
QY 4309 GGGCTGTGCGCTTTAAAGTCAATGAGCGCGAGATGCCCTCCACCGAGCACTGTGTTAAC 4368  
DB 4321 GGGCTGTGCGCTTTAAAGTCAATGAGCGCGAGTGTGCCCTCCACCGAGCACTGTGTTAAC 4380  
QY 4369 CTACTCCCTGTCTATCCTCTCCCTGGCGCCCTAGTGTGTCGGGTGTGTGCGCAGCGATA 4428  
DB 4381 CTACTCCCTGTCTATCCTCTCCCTGGCGCCCTAGTGTGTCGGGTGTGTGCGCAGCGATA 4440  
QY 4429 CTGCGTCGCGACGTGGGCCAGGGAGGGGGCTGTGCACTGATGAAACGGCTGTAGCG 4488  
DB 4441 CTGCGTCGCGACGTGGGCCAGGGAGGGGGCTGTGCACTGATGAAACGGCTGTAGCG 4500  
QY 4489 TTGCGTTTCGGGGTAAACACAGTCTCCGCCACGCACTATGTGCTGAGAGCGACGTGCA 4548  
DB 4501 TTGCGTTTCGGGGTAAACACAGTCTCCGCCACGCACTATGTGCTGAGAGCGACGTGCA 4560  
QY 4549 GCACTGTCACTCAGATCCTCTCTAGTCTTACCACTCACTCAGTGTGAAGAGGCTTCA 4608  
DB 4561 GCACTGTCACTCAGATCCTCTCTAGTCTTACCACTCACTCAGTGTGAAGAGGCTTCA 4620  
QY 4609 CAGTGGATCAACGAGGACTGTCTCCAGCCATGCTCGGCTGTGCTGCTGAGATGTTGG 4668  
DB 4621 CAGTGGATCAACGAGGACTGTCTCCAGCCATGCTCGGCTGTGCTGCTGAGATGTTGG 4680  
QY 4669 GATTGGATGACAGGTTGACTGATTTCAAGACTGCTCCAGTCCAGTCTCTGCGG 4728  
DB 4681 GATTGGATGACAGGTTGACTGATTTCAAGACTGCTCCAGTCCAGTCTCTGCGG 4740  
QY 4729 GATTGGAGTCCCGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4788  
DB 4741 GATTGGAGTCCCGGAGTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4800  
QY 4789 GAGCGATCATGCAAAACCTGCCCCATGTGAGCAGACATCACCAGGACATGTGAAAAC 4848  
DB 4801 GAGCGATCATGCAAAACCTGCCCCATGTGAGCAGACATCACCAGGACATGTGAAAAC 4860  
QY 4849 GGTTCATGAGGATCGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTCCGC 4908  
DB 4861 GGTTCATGAGGATCGTGGGGCTTAGGACCTGTAGTAAACAGTGGCATGGAACATTCCGC 4920  
QY 4909 ATTAAACGGGTACACCAAGCGGCCCTGCAAGCCCTCCCGGCCCAAAATTATTCTAGGGC 4968  
DB 4921 ATTAAACGGGTACACCAAGCGGCCCTGCAAGCCCTCCCGGCCCAAAATTATTCTAGGGC 4980  
QY 4969 CTGTCGGGTGCTGCTGAGGAGTACGTGGAGTTACGCGGGTGGGGATTTCCACTAC 5028  
DB 4981 CTGTCGGGTGAGTGTGAGAGTACGTGGAGTTACGCGGGTGGGGATTTCCACTAC 5040  
QY 5029 GTGACGGGATGACCACTGACAACTGAAAGTCCCGTGTGAGTTCCGGCCCGCAATTC 5088  
DB 5041 GTGACGGGATGACCACTGACAACTGAAAGTCCCGTGTGAGTTCCGGCCCGCAATTC 5100  
QY 5089 TTCAAGAGTGGTGGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 5148  
DB 5101 TTCAAGAGTGGTGGGTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 5160  
QY 5149 CGGAGGAGGTACATCTCTGTCGGGCTCAATCAATCAATCAATCAATCAATCAATCAAT 5208  
DB 5161 CGGAGGAGGTACATCTCTGTCGGGCTCAATCAATCAATCAATCAATCAATCAATCAAT 5220

QY 5209 TCGAGCGCCGAACCGAGCTAGCTGCTCACTTCATGCTCACCGACCCCTCCACATT 5268  
DB 5221 TCGAGCGCTGAACCGGATGAGCTGCTCACTTCATGCTCACCGACCCCTCCACATT 5280  
QY 5269 ACGGGGAGACCGGCTAAGCGTAGGCTGGCCAGGGATCTCCGCCCTCTCTGGCCAGCTCA 5328  
DB 5281 ACGGGGAGACCGGCTAAGCGTAGGCTGGCCAGGGATCTCCGCCCTCTCTGGCCAGCTCA 5340  
QY 5329 TCAGCTAGCAGCTGTCTGCGCTTCTTTGAAGGCAACATGCACTACCCGTATGATCC 5388  
DB 5341 TCAGCTAGCAGCTGTCTGCGCTTCTTTGAAGGCAACATGCACTACCCGTATGATCC 5400  
QY 5389 CGGACGCTGACCTCATCGAGCCCAACTCTCTGTGCGGAGAGATGCGCGGAACTC 5448  
DB 5401 CGGACGCTGACCTCATCGAGCCCAACTCTCTGTGCGGAGAGATGCGCGGAACTC 5460  
QY 5449 ACCCGGTGGAGTCAGAAATAAGGTAGTAAATTTTGGACTCTTTGAGCGCTCCAAAGCG 5508  
DB 5461 ACCCGGTGGAGTCAGAAATAAGGTAGTAAATTTTGGACTCTTTGAGCGCTCCAAAGCG 5520  
QY 5509 GAGGAGGATGAGAGGAGATATCCGTTCCGGCGGAGATCTCGGAGGTCAGGAAATTC 5568  
DB 5521 GAGGAGGATGAGAGGAGATATCCGTTCCGGCGGAGATCTCGGAGGTCAGGAAATTC 5580  
QY 5569 CCTCGAGCGATGCCCATATGGGACGCGCGGATTACACCCCTCACTGTTAGAGTCTCTG 5628  
DB 5581 CCTCGAGCGATGCCCATATGGGACGCGCGGATTACACCCCTCACTGTTAGAGTCTCTG 5640  
QY 5629 AAGGACCCGACTACGTCCTCTCAGTGTGATACCGGCTTCCATTCGCGCTCCCAAGGCC 5688  
DB 5641 AAGGACCCGACTACGTCCTCTCAGTGTGATACCGGCTTCCATTCGCGCTCCCAAGGCC 5700  
QY 5689 CCTCGATACCACTCTCCAGGAGGAGGAGGCTGTCTCTGATCAGATCTACCGTCT 5748  
DB 5701 CCTCGATACCACTCTCCAGGAGGAGGAGGCTGTCTCTGATCAGATCTACCGTCT 5760  
QY 5749 TCTGCTTTGGCGAGCTCGCCAAAGACCTTTGGCAGCTCCGAATCTGTCGCGCTCGAC 5808  
DB 5761 TCTGCTTTGGCGAGCTCGCCAAAGACCTTTGGCAGCTCCGAATCTGTCGCGCTCGAC 5820  
QY 5809 AGCGGACCGGAAAGGCTCTCTGACAGCCCTCCGACGAGCGGAGCGCGGATCCGAC 5868  
DB 5821 AGCGGACCGGAAAGGCTCTCTGACAGCCCTCCGACGAGCGGAGCGCGGATCCGAC 5880  
QY 5869 GTTAGTGTGACTCTCTGATGCCCCCTCTGAGGGGAGCGCGGAGTCCGATCTCAGC 5928  
DB 5881 GTTAGTGTGACTCTCTGATGCCCCCTCTGAGGGGAGCGCGGAGTCCGATCTCAGC 5940  
QY 5929 GACGGGTCTTGGTCTACCGTAAGCGAGGAGGCTAGTGAGGAGCTGCTGCTGCTCGATG 5988  
DB 5941 GACGGGTCTTGGTCTACCGTAAGCGAGGAGGCTAGTGAGGAGCTGCTGCTGCTCGATG 6000  
QY 5989 TCCTACATGAGACAGGCGCTGATCAGCCATCGCTGCGGAGGAGAAACCAAGTGGCC 6048  
DB 6001 TCCTACATGAGACAGGCGCTGATCAGCCATCGCTGCGGAGGAGAAACCAAGTGGCC 6060  
QY 6049 ATCAATGACTGAGCAACTCTTTGCTCCGTACCACTGCTGCTGCTGCTGCTGCTGCTGCT 6108  
DB 6061 ATCAATGACTGAGCAACTCTTTGCTCCGTACCACTGCTGCTGCTGCTGCTGCTGCTGCT 6120  
QY 6109 CGAGCGCAAGCTCGCGGAGAGGAGGTCACCTTTGACAGACTGCACTGCTGCGACGAC 6168  
DB 6121 CGAGCGCAAGCTCGCGGAGAGGAGGTCACCTTTGACAGACTGCACTGCTGCGACGAC 6180  
QY 6169 CACTACCGGAGCTGCTCAAGGAGATGAAGGAGGCTCCACAGTTAAGGCTAACT 6228  
DB 6181 CACTACCGGAGCTGCTCAAGGAGATGAAGGAGGCTCCACAGTTAAGGCTAACT 6240  
QY 6229 CTATCCGTTGAGGAGGCTGTAAGCTGACGCCCACTGCGGAGATCTAAATTTGGC 6288  
DB 6241 CTATCCGTTGAGGAGGCTGTAAGCTGACGCCCACTGCGGAGATCTAAATTTGGC 6300  
QY 6289 TATGGGCAAGGAGCTCCGGAACCTATCCAGCAAGGCGGTTAAACCATCGCTCCGCTG 6348

Db 6301 TATGGGCAAGAGAGCTCCGGAACCTTACAGCAAGCCGCTTACACATCCGCTCCGTG 6360  
Qy 6349 TGAAGGACTTGTGGAAGACACTGAGACACCAATGACACCAATCATATGGCAAAAAAT 6408  
Db 6361 TGAAGGACTTGTGGAAGACACTGAGACACCAATGACACCAATCATATGGCAAAAAAT 6420  
Qy 6409 GAGGTTTCTGCGTCCAAACAGAGAAAGGGGGCGCAAGCCAGCTCGCTTATCGTATTC 6468  
Db 6421 GAGGTTTCTGCGTCCAAACAGAGAAAGGGGGCGCAAGCCAGCTCGCTTATCGTATTC 6480  
Qy 6469 CCAGATTGGGGGTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTC 6528  
Db 6481 CCAGATTGGGGGTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCCCTC 6540  
Qy 6529 CCTCAGCGCGTATGGGCTCTTTCATACGAGTCCAAATCTCTCTGACAGCGGTCGAG 6588  
Db 6541 CCTCAGCGCGTATGGGCTCTTTCATACGAGTCCAAATCTCTCTGACAGCGGTCGAG 6600  
Qy 6589 TTCCTGTGTAATGCTCGGAAGCGAAGAAATGCCCTATGGCTTCGCATATGACACCGC 6648  
Db 6601 TTCCTGTGTAATGCTCGGAAGCGAAGAAATGCCCTATGGCTTCGCATATGACACCGC 6660  
Qy 6649 TGTCTTGACTCAACCGGTCACTGAGAATGACATCCGTTGTGAGGAGTCAATCTACCAATGT 6708  
Db 6661 TGTCTTGACTCAACCGGTCACTGAGAATGACATCCGTTGTGAGGAGTCAATCTACCAATGT 6720  
Qy 6709 TGTGACTTGGCCCCCGAAGCCAGACAGGCCATAAGTTCGCTCAAGAGCGGCTTTTACATC 6768  
Db 6721 TGTGACTTGGCCCCCGAAGCCAGACAGGCCATAAGTTCGCTCAAGAGCGGCTTTTACATC 6780  
Qy 6769 GGGGGCCCCCTGACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGC 6828  
Db 6781 GGGGGCCCCCTGACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGCGAGC 6840  
Qy 6829 GGTGTACTGACGACGAGCTGGGTAAATACCTTCACATGTACTTGAAGCGCGCTGCGGCC 6888  
Db 6841 GGTGTACTGACGACGAGCTGGGTAAATACCTTCACATGTACTTGAAGCGCGCTGCGGCC 6900  
Qy 6889 TGTGAGCTGCGAAGCTCCAGGACTGACGAGTTCGCTATGCGGAGACGACCTTGTGCTT 6948  
Db 6901 TGTGAGCTGCGAAGCTCCAGGACTGACGAGTTCGCTATGCGGAGACGACCTTGTGCTT 6960  
Qy 6949 ATCTGTGAAGCGGGGAGCCAGAGACGAGCGAGCCCTACGGGCTTCAAGGAGCT 7008  
Db 6961 ATCTGTGAAGCGGGGAGCCAGAGAGAGAGCGAGCGCCCTTCAAGGAGCT 7020  
Qy 7009 ATGACTAGATCTGCCCCCCTGGGAGCCCGCCCAACAGAAATACGACTTGGAGTTG 7068  
Db 7021 ATGACTAGATCTGCCCCCCTGGGAGCCCGCCCAACAGAAATACGACTTGGAGTTG 7080  
Qy 7069 ATAAATCATGCTCCTCAATGTGTAGTGTGCGGACGATGTCGCAAAAGGTTGATC 7128  
Db 7081 ATAAATCATGCTCCTCAATGTGTAGTGTGCGGACGATGTCGCAAAAGGTTGATC 7140  
Qy 7129 TATCTACCCGTCACCCACCCCTTGGCGGGCTGCGTGGGAGACGCTAGACAC 7188  
Db 7141 TATCTACCCGTCACCCACCCCTTGGCGGGCTGCGTGGGAGACGCTAGACAC 7200  
Qy 7189 ACTCCAGTCAANTCTGGCTAGGCAACATCATATGTATGCGCCACCTTGTGGGCAAGG 7248  
Db 7201 ACTCCAGTCAANTCTGGCTAGGCAACATCATATGTATGCGCCACCTTGTGGGCAAGG 7260  
Qy 7249 ATGATCTGATGATCATTTCTTCTCCATGCTTCTAGCTCAGGAACTTGAAGAAAGCC 7308  
Db 7261 ATGATCTGATGATCATTTCTTCTCCATGCTTCTAGCTCAGGAACTTGAAGAAAGCC 7320  
Qy 7309 CTAGATTGTGATCTAGCGGGCTGTACTTCCATTGAGCCATTGACCTACCTCAGATC 7368  
Db 7321 CTAGATTGTGATCTAGCGGGCTGTACTTCCATTGAGCCATTGACCTACCTCAGATC 7380  
Qy 7369 ATTCAAGACTCCATGGCCCTTAGCGCAATTTTCACTCCATGTTACTCTCAGGTTGATC 7428

Db 7381 ATTCAAGACTCCATGGCCCTTAGCGCAATTTTCACTCCATAGTTACTCTCAGGTTGATC 7440  
Qy 7429 AATAGGTTGGCTTCATGCTCAGGAACTTTGGGGTACCGCCCTTGCAGTCTGAGACAT 7488  
Db 7441 AATAGGTTGGCTTCATGCTCAGGAACTTTGGGGTACCGCCCTTGCAGTCTGAGACAT 7500  
Qy 7489 CGGGCCAGAAAGTGTCCGCGTGTAGCTACTGTCCAGGGGGGGGGTGCACACTTGTGGC 7548  
Db 7501 CGGGCCAGAAAGTGTCCGCGTGTAGCTACTGTCCAGGGGGGGGGTGCACACTTGTGGC 7560  
Qy 7549 AAGTACCTCTTCAACTGGGAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7608  
Db 7561 AAGTACCTCTTCAACTGGGAGTAAGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCG 7620  
Qy 7609 TCCAGTTGGAATTTATCCAGCTGTTCTGTGCTTGTTCAGAGCGGGGAGACATATATC 7668  
Db 7621 TCCAGTTGGAATTTATCCAGCTGTTCTGTGCTTGTTCAGAGCGGGGAGACATATATC 7680  
Qy 7669 AGCTGTCTGTGCGCCGACCCCGCTGTTTCATGTGGTGCCTACTCTCTCTTCTGTAGG 7728  
Db 7681 AGCTGTCTGTGCGCCGACCCCGCTGTTTCATGTGGTGCCTACTCTCTCTTCTGTAGG 7740  
Qy 7729 GTAGGCACTATCTACTCTCCCAACCGGATGAACGGGACCTAAACACCTCCAGGCCAATAGG 7788  
Db 7741 GTAGGCACTATCTACTCTCCCAACCGGATGAACGGGAGCTAAACACCTCCAGGCCAATAGG 7800  
Qy 7789 CCATCTCTTTTTTCCCTTT 7848  
Db 7801 CCATCTCTTTTTTCCCTTT 7860  
Qy 7849 TTTTCTCTTTTTTTCCTCTTT 7908  
Db 7861 TTTTCTCTTTTTTTCCTCTTT 7920  
Qy 7909 CCTAGTCACGGCTAGCTGTGAAGGTCGAGCGGCTTCACTGCAGAGAGTGTCTGATC 7968  
Db 7921 CCTAGTCACGGCTAGCTGTGAAGGTCGAGCGGCTTCACTGCAGAGAGTGTCTGATC 7980  
Qy 7969 TGGCTCTCTGCAGATCAAGT 7989  
Db 7981 TGGCTCTCTGCAGATCAAGT 8001

RESULT 6  
US-09-539-601-4  
; Sequence 4, Application US/09539601C  
; Patent No. 6630343  
; GENERAL INFORMATION:  
; APPLICANT: Bartenschlager, Ralf FW  
; TITLE OF INVENTION: Hepatitis C Virus Cell Culture System  
; FILE REFERENCE: all sequences  
; CURRENT APPLICATION NUMBER: US/09/539,601C  
; EARLIER FILING DATE: 2001-08-30  
; EARLIER APPLICATION NUMBER: 199 15 178.4 GERMANY  
; EARLIER FILING DATE: 1999-04-03  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 8637  
; TYPE: DNA  
; ORGANISM: Hepatitis C virus  
; FEATURE:  
; NAME/KEY: 5'UTR  
; LOCATION: (1)..(341)  
; OTHER INFORMATION: construct I377/NS2-3'/wt  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (342)..(1181)  
; OTHER INFORMATION: HCV core-neomycin phosphotransferase fusion  
; OTHER INFORMATION: protein  
; FEATURE:  
; NAME/KEY: RBS  
; LOCATION: (1190)..(1800)

OTHER INFORMATION: internal ribosome entry site from  
OTHER INFORMATION: encephalomyocarditis virus

## FEATURE:

NAME/KEY: CDS

LOCATION: (1801)..(8406)

OTHER INFORMATION: hepatitis C virus NS2 - 5B

## FEATURE:

NAME/KEY: 3'UTR

LOCATION: (8407)..(8637)

PUBLICATION INFORMATION:

AUTHORS: Lohmann, Volker

AUTHORS: Krner, Frank

AUTHORS: Koch, Jan-Oliver

AUTHORS: Hexian, Ulrike

AUTHORS: Theilmann, Lorenz

AUTHORS: Bartenschlager, Ralf

TITLE: Replication of subgenomic hepatitis c virus RNAs in a

TITLE: hepatoma cell line

JOURNAL: Science

VOLUME: 285

PAGES: 110-113

DATE: 1999-07-02

US-09-539-601-4

Query Match 91.7%; Score 7329.4; DB 4; Length 8637;

Best Local Similarity 92.5%; Pred. No. 0;

Matches 7988; Conservative 0; Mismatches 1; Indels 648; Gaps 1;

QY	1	GCCAGCCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
DB	1	GCCAGCCCCCGATTGGGGCGGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
QY	61	TCTTACGAGAGAGCGCTAGCCATGGCGTTAGTATGAGTGTCTGCGAGCCCTCCAGGAC 120
DB	61	TCTTACGAGAGAGCGCTAGCCATGGCGTTAGTATGAGTGTCTGCGAGCCCTCCAGGAC 120
QY	121	CCCCCTCCCGGAGAGCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAATTCGCAAG 180
DB	121	CCCCCTCCCGGAGAGCATAGTGTCTGCGGAAACCGGTGAGTACACCGGAATTCGCAAG 180
QY	181	GACGACCGGGTCTTTCTTGGATCAACCGCTCAATGCTGAGGATTTGGGGGTGCCCCC 240
DB	181	GACGACCGGGTCTTTCTTGGATCAACCGCTCAATGCTGAGGATTTGGGGGTGCCCCC 240
QY	241	GGAGACTCTAGCCGAGTAGTGTGGGTGCGGAAAGGGCTTTGTGGTACTGCGCTGATAGG 300
DB	241	GGAGACTCTAGCCGAGTAGTGTGGGTGCGGAAAGGGCTTTGTGGTACTGCGCTGATAGG 300
QY	301	GTGCTTGGAGTGCCCGGGAGGTCTCTGAGACCGGTGACCATGAGCAGCAATCTCTAAAC 360
DB	301	GTGCTTGGAGTGCCCGGGAGGTCTCTGAGACCGGTGACCATGAGCAGCAATCTCTAAAC 360
QY	361	CTCAAGAAAAACCAAGAGCGCGCATGATTGAACAAGATGGATTCGACGAGGTTCTC 420
DB	361	CTCAAGAAAAACCAAGAGCGCGCATGATTGAACAAGATGGATTCGACGAGGTTCTC 420
QY	421	CGGCGCTTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAACAGCAATCGGCTGCT 480
DB	421	CGGCGCTTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAACAGCAATCGGCTGCT 480
QY	481	CTGATGCGCGCTTTCCGGCTGTACGCGAGGGGCGCGGTTCTTTTGTCAAGACCG 540
DB	481	CTGATGCGCGCTTTCCGGCTGTACGCGAGGGGCGCGGTTCTTTTGTCAAGACCG 540
QY	541	ACCTGTCCGGTCCCTTGAATGAATCTGAGGAGCGAGCGGCTATCGTGGCTGGCCA 600
DB	541	ACCTGTCCGGTCCCTTGAATGAATCTGAGGAGCGAGCGGCTATCGTGGCTGGCCA 600
QY	601	CGACGGCGCTTCTTCCGAGCTGTCTGAGCGTGTCTGAGCGTGTCTGAGCGGAGGAGCTGCG 660
DB	601	CGACGGCGCTTCTTCCGAGCTGTCTGAGCGTGTCTGAGCGTGTCTGAGCGGAGGAGCTGCG 660
QY	661	TGCTATTGGCGAAGTGCCGGGAGGATCTCTGTCTCATCTCACCTTGCTGCGGAGA 720

DB	661	TGCTATTGGCGAAGTGCCGGGAGGATCTCTGTCTCATCTCACCTTGCTGCGGAGA 720
QY	721	AAGTATCCATCATGCTGATGCAATGCGCGGCTGATGCTGATCGGCTACTGCTGC 780
DB	721	AAGTATCCATCATGCTGATGCAATGCGCGGCTGATGCTGATCGGCTACTGCTGC 780
QY	781	CATTGCAACCAAGCGAATCCATCGAGGAGCAGTACTCGGATGGAAGCGGTC 840
DB	781	CATTGCAACCAAGCGAATCCATCGAGGAGCAGTACTCGGATGGAAGCGGTC 840
QY	841	TTGTCCATCAGGATGATCTGGAAGAGCAGTACAGGGCTCGCGCAGCGAATCTTCG 900
DB	841	TTGTCCATCAGGATGATCTGGAAGAGCAGTACAGGGCTCGCGCAGCGAATCTTCG 900
QY	901	CCAGGCTCAAGCGCGCATGCCCCGAGGAGTCTGCTGAGCCATGCGGATGCTT 960
DB	901	CCAGGCTCAAGCGCGCATGCCCCGAGGAGTCTGCTGAGCCATGCGGATGCTT 960
QY	961	GCTTCCGGAATATCATGTTGGAATGCGCGCTTTTCTGATTCATCGACTGTGCGCGC 1020
DB	961	GCTTCCGGAATATCATGTTGGAATGCGCGCTTTTCTGATTCATCGACTGTGCGCGC 1020
QY	1021	TGGGTGTGGCGACCGCTATCAGGACATAGCCGTGCTACCGCTGATATTGCTGAAGAGC 1080
DB	1021	TGGGTGTGGCGACCGCTATCAGGACATAGCCGTGCTACCGCTGATATTGCTGAAGAGC 1080
QY	1081	TTGGCGGCGAATGCGGTGACCGCTTCTGCTGCTTTTACCGGTATCGCGCTCCGATTCGC 1140
DB	1081	TTGGCGGCGAATGCGGTGACCGCTTCTGCTGCTTTTACCGGTATCGCGCTCCGATTCGC 1140
QY	1141	AGCGCATCGCTTCTATCGCTTCTGAGAGTCTCTTCTGAGTTTAAACAGACCAACG 1200
DB	1141	AGCGCATCGCTTCTATCGCTTCTGAGAGTCTCTTCTGAGTTTAAACAGACCAACG 1200
QY	1201	GTTCCTCTTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAACTGCTG 1260
DB	1201	GTTCCTCTTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTAACTGCTG 1260
QY	1261	CGAAGCGCTTGAATAGCGCGGTGCTGCTTCTATATGTTATTTCCACCATTTG 1320
DB	1261	CGAAGCGCTTGAATAGCGCGGTGCTGCTTCTATATGTTATTTCCACCATTTG 1320
QY	1321	CGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCTGCTTCTTGAAGAGCATTCCT 1380
DB	1321	CGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCTGCTTCTTGAAGAGCATTCCT 1380
QY	1381	AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGCTCGTGAAGAGCA 1440
DB	1381	AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGCTCGTGAAGAGCA 1440
QY	1441	GTTCTCTGGAAGCTTTTGAAGACAAACAGTCTGTGAGGACCCCTTTGAGGAGCGG 1500
DB	1441	GTTCTCTGGAAGCTTTTGAAGACAAACAGTCTGTGAGGACCCCTTTGAGGAGCGG 1500
QY	1501	AAACCCCGCTGCGGACAGGTGCTTCTGCGGCAAAAGCCACGTGATATAAGATACACCT 1560
DB	1501	AAACCCCGCTGCGGACAGGTGCTTCTGCGGCAAAAGCCACGTGATATAAGATACACCT 1560
QY	1561	GCAAGGCGGCAACCCCGAGTGCACAGTGTGAGTTGATAGTTGTGGAAGAGTCAAA 1620
DB	1561	GCAAGGCGGCAACCCCGAGTGCACAGTGTGAGTTGATAGTTGTGGAAGAGTCAAA 1620
QY	1621	TGGCTCTCTAAGGCTATTCAACAGGGGCTGAGGATGCCAGAGGTACCCCATTTGT 1680
DB	1621	TGGCTCTCTAAGGCTATTCAACAGGGGCTGAGGATGCCAGAGGTACCCCATTTGT 1680
QY	1681	ATGGGATCTGATCTGGGCTCGGTGACATGCTTTTACATGCTTTTGTGAGTGAAGTAA 1740
DB	1681	ATGGGATCTGATCTGGGCTCGGTGACATGCTTTTACATGCTTTTGTGAGTGAAGTAA 1740
QY	1741	AACGCTTAGGCGCCCGGAAACCGGAGCGTGGTTTCTTTGAAAAACAGGATATACC 1800



Db 1741 AACGTCTAGGCCCCCGAACCACCGGAGCGTGTTCCTTTGAAABACACGATAATACC 1800  
QY 1801 ATG----- 1803  
Db 1801 ATGGACGGGAGATGGCAGCATCGTGGAGGCGGTTTTTCGTAGGTCTGATACTCTTG 1860  
QY 1804 ----- 1803  
Db 1861 ACCTTGTACCGCACTATAAGCTGTTCCTCGCTAGGCTCATATGGTGGTTACAATATTTT 1920  
QY 1804 ----- 1803  
Db 1921 ATCACCAGGCGGAGGCACACTTGAAGTGTGGATCCGCCCTCAAGTTTCGGGGGGC 1980  
QY 1804 ----- 1803  
Db 1981 CGGATGCCGTATCTCTCTCACTGCGCGATCCACCAGAGCTAATCTTTTACCATCACC 2040  
QY 1804 ----- 1803  
Db 2041 AATACTTGTCTCGCATACTCGGTCCACTCATGTGTCTCAGGCTGGTATTAACCAAGTG 2100  
QY 1804 ----- 1803  
Db 2101 CCGTACTTCTGGCGCACACGGGCTCATTCGTGATGATCTGTGCGGAAGTTGCT 2160  
QY 1804 ----- 1803  
Db 2161 GGGGTCAATTATGTCCAAATGGCTCTCATGAAGTTGGCCGCACTGACAGGTACGTAGCTT 2220  
QY 1804 ----- 1803  
Db 2221 TATGACCATCTACCCCACTGCGGCACTGGGCCACCGGGCTTACGAGACCTTGGGGTG 2280  
QY 1804 ----- 1803  
Db 2281 GCAGTTGAGCCCGTCTTCTCTGATATGAGACCAAGTTATACCTTGGGGGCGAGC 2340  
QY 1804 ----- 1803  
Db 2341 ACCCGCGGTGTGGGACATCATTTGGGCTGCCCCGCTCTCGGCCGAGGGGAGGAG 2400  
QY 1804 ----- 1812  
Db 2401 ATACATCTGGGACCGGCAGACAGCTTGAAGGCGAGGGGTGGGACCTCTCGCGCCCTATT 2460  
QY 1813 ACGGCTTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCACTAGCCTTCAAGGC 1872  
Db 2461 ACGGCTTACTCCCAACAGACGCGAGGCTTACTTGGCTGCATCATCACTAGCCTTCAAGGC 2520  
QY 1873 CGGACAGGAAACAGGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCAACACAAATCTTTC 1932  
Db 2521 CGGACAGGAAACAGGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCAACACAAATCTTTC 2580  
QY 1933 CTGGCAGCTTGCATTAATGGCTGTGTGTGACTGTCTATCATGTGTCGGCTCAAGAGCC 1992  
Db 2581 CTGGCAGCTTGCATTAATGGCTGTGTGTGACTGTCTATCATGTGTCGGCTCAAGAGCC 2640  
QY 1993 CTTCGGGCGCCAAAGGCGCCAAATCACCCAAATGTACACCAATGTGACACAGGACCTCTGTC 2052  
Db 2641 CTTCGGGCGCCAAAGGCGCCAAATCACCCAAATGTACACCAATGTGACACAGGACCTCTGTC 2700  
QY 2053 GGTGCAAGGCGCCCGCGGGCGGTTCTTGTACACCATGTGACCTGTGGCGAGCTGGGAC 2112  
Db 2701 GGTGCAAGGCGCCCGCGGGCGGTTCTTGTACACCATGTGACACCATGTGACCTGTGGCGAGCTGGGAC 2760  
QY 2113 CTTTACTTGTTCAGAGGCATGCCGATGTTCATTCGGGTGCGCGGGCGGCGACAGG 2172  
Db 2761 CTTTACTTGTTCAGAGGCATGCCGATGTTCATTCGGGTGCGCGGGCGGCGACAGG 2820  
QY 2173 GGGAGCCTACTCTCCCCCAGGCGGCTCTCCTACTTGAAGGGCTCTTCGGGCGGTCCACTG 2232  
Db 2821 GGGAGCCTACTCTCCCCCAGGCGGCTCTCCTACTTGAAGGGCTCTTCGGGCGGTCCACTG 2880

QY 2233 CTCTGCCCCCTCGGCGACGCTGTGGGCATCTTTTCGGGCTGCCGTGTGACCCGAGGGTT 2292  
Db 2881 CTCTGCCCCCTCGGCGACGCTGTGGGCATCTTTTCGGGCTGCCGTGTGACCCGAGGGTT 2940  
QY 2293 GCGAAGCGGTTGACTTTGTACCGTTCGAGTCTATGGAACCACTATGCGGTCCCGGTC 2352  
Db 2941 GCGAAGCGGTTGACTTTGTACCGTTCGAGTCTATGGAACCACTATGCGGTCCCGGTC 3000  
QY 2353 TTACGAGCAACTCGTCCCTCCCGCGGTACCGCAGACATTTCCAGGTGGCCCATCTACAC 2412  
Db 3001 TTACGAGCAACTCGTCCCTCCCGCGGTACCGCAGACATTTCCAGGTGGCCCATCTACAC 3060  
QY 2413 GCCCCTACTGCTAGCGGCAAGACACTAAAGTGGCGCTGCTATGAGGCCAAGGTTAT 2472  
Db 3061 GCCCCTACTGCTAGCGGCAAGACACTAAAGTGGCGCTGCTATGAGGCCAAGGTTAT 3120  
QY 2473 AAGTGTCTTGTCTGAACCGTTCGTCGCGCCACCTAGGTTTCGGGGGCTATATGCT 2532  
Db 3121 AAGTGTCTTGTCTGAACCGTTCGTCGCGCCACCTAGGTTTCGGGGGCTATATGCT 3180  
QY 2533 AAGSCACATGCTATCGACCCCTAACATCAGAACCGGGGTAAAGGACCATCAACACGGGTGCC 2592  
Db 3181 AAGSCACATGCTATCGACCCCTAACATCAGAACCGGGGTAAAGGACCATCAACACGGGTGCC 3240  
QY 2593 CCATCAGTACTCCACCTATGGCAAGTTTCTTGGCGACGGTGGTGTCTGTGGGGGCGCC 2652  
Db 3241 CCATCAGTACTCCACCTATGGCAAGTTTCTTGGCGACGGTGGTGTCTGTGGGGGCGCC 3300  
QY 2653 TATGACATCATATATGTGATGATGCCACTCAACTGACCTCGACCATCTCTTGGGATC 2712  
Db 3301 TATGACATCATATATGTGATGATGCCACTCAACTGACCTCGACCATCTCTTGGGATC 3360  
QY 2713 GGCACAGTCTCGACCAAGCGGAGACGGCTGGAGCGGACTCGTCTGCTCGCCACCGCT 2772  
Db 3361 GGCACAGTCTCGACCAAGCGGAGACGGCTGGAGCGGACTCGTCTGCTCGCCACCGCT 3420  
QY 2773 ACGCTCCGGATCGGTACCGTGCCACATCCAAAATCGAGGAGTGGTCTGTCTCAGC 2832  
Db 3421 ACGCTCCGGATCGGTACCGTGCCACATCCAAAATCGAGGAGTGGTCTGTCTCAGC 3480  
QY 2833 ACTGGAGAAATCCCTTTTATGCAAGCCATCCCATCGAGACCATCAAGGGGGGAGG 2892  
Db 3481 ACTGGAGAAATCCCTTTTATGCAAGCCATCCCATCGAGACCATCAAGGGGGGAGG 3540  
QY 2893 CACTCATTTTCTGCCATTCGAAGAAATGTGATGAGCTCGCCCGGAAGCTGTCCGGC 2952  
Db 3541 CACTCATTTTCTGCCATTCGAAGAAATGTGATGAGCTCGCCCGGAAGCTGTCCGGC 3600  
QY 2953 CTGGACTCAATGCTGTAGCATATTAACCGGGGCTTGAATGATATCCGTCATACCAACTAGC 3012  
Db 3601 CTGGACTCAATGCTGTAGCATATTAACCGGGGCTTGAATGATATCCGTCATACCAACTAGC 3660  
QY 3013 GGAGAGCTATTCTCGTACGACGCGCTCTAATGACGGGCTTACCGCGGATTCGAC 3072  
Db 3661 GGAGAGCTATTCTCGTACGACGCGCTCTAATGACGGGCTTACCGCGGATTCGAC 3720  
QY 3073 TCAGTGATCGACTGCAATACATGTGTCAACGACAGTTCGACTTCAGCTTGGACCCGAGC 3132  
Db 3721 TCAGTGATCGACTGCAATACATGTGTCAACGACAGTTCGACTTCAGCTTGGACCCGAGC 3780  
QY 3133 TTCAACATTTAGACAGCAGCGGTGCCAAGACGCGGTGTCAAGCTTCGACGCGCGGAGGC 3192  
Db 3781 TTCAACATTTAGACAGCAGCGGTGCCAAGACGCGGTGTCAAGCTTCGACGCGCGGAGGC 3840  
QY 3193 AGGACTGGTAGGGGAGGATGGGCATTTACAGTTTGTCACTCCAGAGAAACGGCCCTCG 3252  
Db 3841 AGGACTGGTAGGGGAGGATGGGCATTTACAGTTTGTCACTCCAGAGAAACGGCCCTCG 3300  
QY 3253 GGCATGCTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGCTGTCTTGGTACGAG 3312  
Db 3901 GGCATGCTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGCTGTCTTGGTACGAG 3960

3313 QY CTCACGCCCGGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACACACAGGGTTGCC 3372  
Db CTCACGCCCGGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACACACAGGGTTGCC 4020  
3373 QY GTCTCCAGGACCATCTGGAGTTCTGGGAGCGCTTTTACAGGCTCAACCCACATAGAC 3432  
Db GTCTCCAGGACCATCTGGAGTTCTGGGAGCGCTTTTACAGGCTCAACCCACATAGAC 4080  
3433 QY GCCCATTTCTGTCCAGACTAAGCAGGAGAGCAACTTCCCTACCTGTAGCATAC 3492  
Db GCCCATTTCTGTCCAGACTAAGCAGGAGAGCAACTTCCCTACCTGTAGCATAC 4140  
3493 QY CAGGCTACGCTGTGCGCAGGCTCAGGCTCCACCTCCATCTGGGAGCAAAATGTGGAAG 3552  
Db CAGGCTACGCTGTGCGCAGGCTCAGGCTCCACCTCCATCTGGGAGCAAAATGTGGAAG 4200  
3553 QY TGCTCATACGGCTAAAGCCTACGCTGCAACGGGCCAAAGCCCTGCTGTATAGGCTGGGA 3612  
Db TGCTCATACGGCTAAAGCCTACGCTGCAACGGGCCAAAGCCCTGCTGTATAGGCTGGGA 4260  
3613 QY GCGGTTCAAAACGAGGTTACTACCAACACACCCCATTAACCAATACATCATGCGATGCA 3672  
Db GCGGTTCAAAACGAGGTTACTACCAACACACCCCATTAACCAATACATCATGCGATGCA 4320  
3673 QY TCGGCTGACCTGAGGTCGTACAGCACTGGGTGCTGGTAGGCGGAGTCCCTAGCAGCT 3732  
Db TCGGCTGACCTGAGGTCGTACAGCACTGGGTGCTGGTAGGCGGAGTCCCTAGCAGCT 4380  
3733 QY CTGGGCGGCTATGCTGCAACAGGACGCGTGGTCATTTGGGCGAGGATCATCTTTGTC 3792  
Db CTGGGCGGCTATGCTGCAACAGGACGCGTGGTCATTTGGGCGAGGATCATCTTTGTC 4440  
3793 QY GGAAGCGCGCATCATCTCCGACAGGAGTCTTTACCGGGAGTTCGATGAGATGGA 3852  
Db GGAAGCGCGCATCATCTCCGACAGGAGTCTTTACCGGGAGTTCGATGAGATGGA 4500  
3853 QY GAGTGGCTCAACCTCCTTACATCGAACAGGAGTGCAGCTCGCCGAAACAAATTCAAA 3912  
Db GAGTGGCTCAACCTCCTTACATCGAACAGGAGTGCAGCTCGCCGAAACAAATTCAAA 4560  
3913 QY CAGAAGGCAATCGGTTGTGCAACAGCCACCAAGCAAGCGGAGGCTGCTGCCGCTG 3972  
Db CAGAAGGCAATCGGTTGTGCAACAGCCACCAAGCAAGCGGAGGCTGCTGCCGCTG 4620  
3973 QY GTGGATCAAGTGGCGGACCTCGAAGCTTCTGGGCGAAGCATATGTGGAATTCATC 4032  
Db GTGGATCAAGTGGCGGACCTCGAAGCTTCTGGGCGAAGCATATGTGGAATTCATC 4680  
4033 QY AGCGGATCAATATTTAGCAGGCTTGTCCACTCTGCTGCTGGCAACCCCGGATAGCATCA 4092  
Db AGCGGATCAATATTTAGCAGGCTTGTCCACTCTGCTGCTGGCAACCCCGGATAGCATCA 4740  
4093 QY CTGATGGCATTCACAGCTCTATCACAGCCGCTCAGCAACCAACATACCTCTCTGTTT 4152  
Db CTGATGGCATTCACAGCTCTATCACAGCCGCTCAGCAACCAACATACCTCTCTGTTT 4800  
4153 QY AACATCTGGGGGATGGGTCGCGCCCACTTGTCTCCAGCGCTGCTCTGCTTTC 4212  
Db AACATCTGGGGGATGGGTCGCGCCCACTTGTCTCCAGCGCTGCTCTGCTTTC 4860  
4213 QY GTAGGCGCGGATCGCTGAGCGGCTGTTGGCAGCATAGGCTTGGGAAGTGTG 4272  
Db GTAGGCGCGGATCGCTGAGCGGCTGTTGGCAGCATAGGCTTGGGAAGTGTG 4920  
4273 QY GATATTTTGGCAGGTTATGAGCAGGCTGGCAGGCGCTCTGGGCTTTAAGGTCATG 4332  
Db GATATTTTGGCAGGTTATGAGCAGGCTGGCAGGCGCTCTGGGCTTTAAGGTCATG 4980  
4333 QY AGCGGCGAGATGCCCTCCACCGAGGACCTGTTTAACTACTCTCTGCTATCTCTCCCT 4392  
Db AGCGGCGAGATGCCCTCCACCGAGGACCTGTTTAACTACTCTCTGCTATCTCTCCCT 5040  
4393 QY GGCGCCCTAGTCTCGGGGTCTGTGCGCAGCATCTGCTCGGCACTGCTGGGCCAGGG 4452

5041 Db GGCGCCCTAGTCTCGGGGTCTGTGCGCAGCATACTCTGCTCGGCACGTTGGGCCACGG 5100  
4453 QY GAGGGGGTGTGCTAGTGAACCGGCTGATAGGTTCTGCTTCGCGGGTAAACACGCTC 4512  
Db GAGGGGGTGTGCTAGTGAACCGGCTGATAGGTTCTGCTTCGCGGGTAAACACGCTC 5160  
4513 QY TCCCCCAACGCACTATGTGCTGAGAGCGACGCTGACGACGCTGTCACTCAGATCTCTCT 4572  
Db TCCCCCAACGCACTATGTGCTGAGAGCGACGCTGACGACGCTGTCACTCAGATCTCTCT 5220  
4573 QY AGTCTTACATCACTCAGCTGTGAAGGCTTCAACGAGTGAATCAACGAGACTGCTTCC 4632  
Db AGTCTTACATCACTCAGCTGTGAAGGCTTCAACGAGTGAATCAACGAGACTGCTTCC 5280  
4633 QY AGCCCATGCTCGGCTCGTGGCTAAGAGATGTTGGGATGATGACAGGTTGACT 4692  
Db AGCCCATGCTCGGCTCGTGGCTAAGAGATGTTGGGATGATGACAGGTTGACT 5340  
4693 QY GATTTCAAGACCTGGCTCCAGTCCAGCTCCTGCCGCGATTCCCGGAGTCCCTTCTTC 4752  
Db GATTTCAAGACCTGGCTCCAGTCCAGTCCAGCTCCTGCCGCGATTCCCGGAGTCCCTTCTTC 5400  
4753 QY TCATGTCAACGTTGGGTAACGAGGAGTCTGGCGGGGCGACGGCATCATGCAAAACCACTGC 4812  
Db TCATGTCAACGTTGGGTAACGAGGAGTCTGGCGGGGCGACGGCATCATGCAAAACCACTGC 5460  
4813 QY CCATGTGAGACACATCACCGACATGTGAACGCTTCCATGAGGATCGTGGGCGCT 4872  
Db CCATGTGAGACACATCACCGACATGTGAACGCTTCCATGAGGATCGTGGGCGCT 5520  
4873 QY AGGACCTGTAGTAACACGCTGGCATGGAACATTTCCCATTAACCGCTTACACCGGCGCC 4932  
Db AGGACCTGTAGTAACACGCTGGCATGGAACATTTCCCATTAACCGCTTACACCGGCGCC 5580  
4933 QY TGACGCGCTTCCCGGCGCAAAATTTATTTAGGCGCTCTGGCGGGTGGCTCTGAGGAG 4992  
Db TGACGCGCTTCCCGGCGCAAAATTTATTTAGGCGCTCTGGCGGGTGGCTCTGAGGAG 5640  
4993 QY TAGCTGGAGTTACGCGGTGGGATTTCCACTACTGACGGCATGACCACTGACAAC 5052  
Db TAGCTGGAGTTACGCGGTGGGATTTCCACTACTGACGGCATGACCACTGACAAC 5700  
5053 QY GTAAAGTGCCTGTGTCAGGTTCCGCGCCCGCAATTTCTTACAGAAAGTGGATGGGTGCGG 5112  
Db GTAAAGTGCCTGTGTCAGGTTCCGCGCCCGCAATTTCTTACAGAAAGTGGATGGGTGCGG 5760  
5113 QY TTGCA CAGGTA CCGTCCAGCTGCAAAACCCCTCTTACGGGAGGAGTCACTTCTGCTTC 5172  
Db TTGCA CAGGTA CCGTCCAGCTGCAAAACCCCTCTTACGGGAGGAGTCACTTCTGCTTC 5820  
5173 QY GGGCTCAATCAATACCTGTTGGGTCAACAGCTCCCATGCGGCCGAAACCGGACGTAGCA 5232  
Db GGGCTCAATCAATACCTGTTGGGTCAACAGCTCCCATGCGGCCGAAACCGGACGTAGCA 5880  
5233 QY GTGCTCACTTCCATGCTCCGACCCCTCCCATTTACCGCGGAGACGGCTAAGCGTAGG 5292  
Db GTGCTCACTTCCATGCTCCGACCCCTCCCATTTACCGCGGAGACGGCTAAGCGTAGG 5940  
5293 QY CTGGCCAGGGATCTCCCGCTCTTGGCCAGCTCATAGCTAGCCAGCTGTCTGCGCT 5352  
Db CTGGCCAGGGATCTCCCGCTCTTGGCCAGCTCATAGCTAGCCAGCTGTCTGCGCT 6000  
5353 QY TCCCTTCAAGGCAACATGCACTACCCGCTCATGCTCCCGGAGCGCTGACCTCATCGAGGCC 5412  
Db TCCCTTCAAGGCAACATGCACTACCCGCTCATGCTCCCGGAGCGCTGACCTCATCGAGGCC 6060  
5413 QY AACCTCTGTGGCGGAGGATGGCGGGAACATCAACCGCGGTGGAGTCAAGAAATAG 5472  
Db AACCTCTGTGGCGGAGGATGGCGGGAACATCAACCGCGGTGGAGTCAAGAAATAG 6120  
5473 QY GTAGTAATTTTGAATCTTTTCGAGCGCTCCAGCGGAGGAGTGAAGGAGTATCC 5532

6121	GTAGTAATTTTGGACTCTTTGAGCCCGCTCCAGCGGAGGAGGATGAGAGGGAGTATCC	618
5533	GTTCGGCGGAGATCTCTCGGAGGTCCAGGAAATTCCTCGAGCGATGCCCATATGGCA	5592
6181	GTTCGGCGGAGATCTCTCGGAGGTCCAGGAAATTCCTCGAGCGATGCCCATATGGCA	6240
5593	CGCCGGATTACAAACCTTCCACTGTTAGAGTCTTGGAAAGACCCGGACTACGTCCTCCA	5652
6241	CGCCGGATTACAAACCTTCCACTGTTAGAGTCTTGGAAAGACCCGGACTACGTCCTCCA	6300
5653	GTGTTACACGGGTGTCATTGCGCGCTGCCAAGGCCCTCGATACACCTCCACGAGG	5712
6301	GTGTTACACGGGTGTCATTGCGCGCTGCCAAGGCCCTCGATACACCTCCACGAGG	6360
5713	AAGAGGACGGTTGTCTGTGTCAGAAATCTACCGTGTCTTCGCTTGGCGGAGCTCGCAC	5772
6361	AAGAGGACGGTTGTCTGTGTCAGAAATCTACCGTGTCTTCGCTTGGCGGAGCTCGCAC	6420
5773	AAGACTTTCGGAGCTCCGAATCGTGGCGGTGACAGCGCACCGGCAACGGGCTCTCCT	5832
6421	AAGACTTTCGGAGCTCCGAATCGTGGCGGTGACAGCGCACCGGCAACGGGCTCTCCT	6480
5833	GACCAGCCCTCCGACGACGGCGGATCCGAGCTTCGAGTCGTACTCTCGATGCC	5892
6481	GACCAGCCCTCCGACGACGGCGGATCCGAGCTTCGAGTCGTACTCTCGATGCC	6540
5893	CCCTTTGAGGGGAGCCGGGGATCCCGATCTCAGCAGCGGTCTTGCTCTACCGTAAGC	5952
6541	CCCTTTGAGGGGAGCCGGGGATCCCGATCTCAGCAGCGGTCTTGCTCTACCGTAAGC	6600
5953	GAGGAGCTAGTGAGGACGTGTCTGTCTCGATGTCCTACATGGAACAGCGCGCTTG	6012
6601	GAGGAGCTAGTGAGGACGTGTCTGTCTCGATGTCCTACATGGAACAGCGCGCTTG	6660
6013	ATCAGCCATTGCGTGGGAGGAAACCAAGCTGCCATCAATGCATGACGAACTCTTG	6072
6661	ATCAGCCATTGCGTGGGAGGAAACCAAGCTGCCATCAATGCATGACGAACTCTTG	6720
6073	CTCCGTCAACCAACTTGGTCTATGTCTAACATCTCGACGCAAGCCCTGGCGAGAAG	6132
6721	CTCCGTCAACCAACTTGGTCTATGTCTAACATCTCGACGCAAGCCCTGGCGAGAAG	6780
6133	AAGTCACTTTGACAGACTGCAAGTCTCGACGACCACTACCGGACGCTGTCAAGGAG	6192
6781	AAGTCACTTTGACAGACTGCAAGTCTCGACGACCACTACCGGACGCTGTCAAGGAG	6840
6193	ATGAAGCGAAGCGTCCACAGTTAGGCTTAACTTCTATCCGTGGAGGAGCTGTAG	6252
6841	ATGAAGCGAAGCGTCCACAGTTAGGCTTAACTTCTATCCGTGGAGGAGCTGTAG	6900
6253	CTCAGCCCCCACTTCGGCCAGATCTAAATTTTGGTATGGGGCAAGGACGTCGGGAAC	6312
6901	CTCAGCCCCCACTTCGGCCAGATCTAAATTTTGGTATGGGGCAAGGACGTCGGGAAC	6960
6313	CTATCCAGCAGCGCTTAACACATCCGCTCCGTGTGGAAGCACTGTCTGGAAGACAT	6372
6961	CTATCCAGCAGCGCTTAACACATCCGCTCCGTGTGGAAGCACTGTCTGGAAGACAT	7020
6373	GAGACCAATTGACACCCACCTCATGGCAAAAATGAGGTTTTCTCGCTCCACACAGAG	6432
7021	GAGACCAATTGACACCCACCTCATGGCAAAAATGAGGTTTTCTCGCTCCACACAGAG	7080
6433	AAGGGGGCGCGAAGCCAGCTCGCCTTATCGTATCCAGATTTGGGGTTCTGTGTGC	6492
7081	AAGGGGGCGCGAAGCCAGCTCGCCTTATCGTATCCAGATTTGGGGTTCTGTGTGC	7140
6493	GAGAAATGGCCCTTTACGATGGTCTCCACCTCCCTCAGCCGCTGATGGGCTCTCA	6552
7141	GAGAAATGGCCCTTTACGATGGTCTCCACCTCCCTCAGCCGCTGATGGGCTCTCA	7200
6553	TACGGATTCCAATACTCTCTGACAGCGGTGAGTTCTGTGTGATGCTGTGGAAGCG	6612
7201	TACGGATTCCAATACTCTCTGACAGCGGTGAGTTCTGTGTGATGCTGTGGAAGCG	7260

QY	6613	AAGAAATGCCCTATTGGCGTTTCGCATATGACACCCGCTGTGTTTTGACTCAACCGGTCACTGAG	6673
Db	7261	AAGAAATGCCCTATTGGCGTTTCGCATATGACACCCGCTGTGTTTTGACTCAACCGGTCACTGAG	7320
QY	6673	AATGACATCCGTTGTGAGGAGTCAATCTACCAATGTTGTCACTTGGCCCCCGGAAGCCAGA	6732
Db	7321	AATGACATCCGTTGTGAGGAGTCAATCTACCAATGTTGTCACTTGGCCCCCGGAAGCCAGA	7380
QY	6733	CAGGCCATAAGGTGCCTCACAGAGCGGCTTTACATCGGGGGCCCCCTGACTAATTTCTAAA	6792
Db	7381	CAGGCCATAAGTTCGCTCAAGAGCGGCTTTACATCGGGGGCCCCCTGACTAATTTCTAAA	7440
QY	6793	GGGAGAACTGCGGCTATGCGCGGTATGCGCGAGCGGTGTACTGACGACCACTCGCGGT	6852
Db	7441	GGGAGAACTGCGGCTATGCGCGGTATGCGCGAGCGGTGTACTGACGACCACTCGCGGT	7500
QY	6853	AATACCCCTCACATGTTTACTTTGAAGGCGCGTGC CGGCTGTTCGAGCTCGGAAGTCCAGGAC	6912
Db	7501	AATACCCCTCACATGTTTACTTTGAAGGCGCGTGC CGGCTGTTCGAGCTCGGAAGTCCAGGAC	7560
QY	6913	TGCACGATGCTCGTATGCGGAGACGACTTGTTCGTTTATCTGTCAAAGCGGGGAGCCCAA	6972
Db	7561	TGCACGATGCTCGTATGCGGAGACGACTTGTTCGTTTATCTGTCAAAGCGGGGAGCCCAA	7620
QY	6973	GAGGACGAGCGAGCCCTACGGGCTTTACGGAGGCTATGACTAGACTACTGTGCCCCCCT	7032
Db	7621	GAGGACGAGCGAGCCCTACGGGCTTTACGGAGGCTATGACTAGACTACTGTGCCCCCCT	7680
QY	7033	GGGAGCCGCGCCAAACAGAAATACGACTTGGAGTTGATTAACATCATCTGCTCTCCATGTG	7092
Db	7681	GGGAGCCGCGCCAAACAGAAATACGACTTGGAGTTGATTAACATCATCTGCTCTCCATGTG	7740
QY	7093	TCAGTTCGCGACGATGCATCTGCGAAAGGGTGTACTTCTCAACCGTGCACCCACCACC	7152
Db	7741	TCAGTTCGCGACGATGCATCTGCGAAAGGGTGTACTTCTCAACCGTGCACCCACCACC	7800
QY	7153	CCCCCTTCGCGGGCTGCGTGGGAGACGAGCTAGACACACTCCAGTCAATTTCTGGGTAGGC	7212
Db	7801	CCCCCTTCGCGGGCTGCGTGGGAGACGAGCTAGACACACTCCAGTCAATTTCTGGGTAGGC	7860
QY	7213	AACATCATATGTTATGCGGCCCACTTGTGGGCAAGGATGATCCTGATGACTCATTTCTTTC	7272
Db	7861	AACATCATATGTTATGCGGCCCACTTGTGGGCAAGGATGATCCTGATGACTCATTTCTTTC	7920
QY	7273	TCCATCCTTCTAGCTCAGGACAACTTGAAAAGCGCTAGATGTGCAGATCTACGGGCGC	7332
Db	7921	TCCATCCTTCTAGCTCAGGACAACTTGAAAAGCGCTAGATGTGCAGATCTACGGGCGC	7980
QY	7333	TGTTACTCCAATTGAGCCACTTGACCTACCTCAGATCAATTCACGACTCCATGCGCCTTAGC	7392
Db	7981	TGTTACTCCAATTGAGCCACTTGACCTACCTCAGATCAATTCACGACTCCATGCGCCTTAGC	8040
QY	7393	GCATTTTCACTCCATAGTTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCATGCCTCAGG	7452
Db	8041	GCATTTTCACTCCATAGTTTACTCTCCAGGTGAGATCAATAGGGTGGCTTCATGCCTCAGG	8100
QY	7453	AAACTTTGGGTTACCGCCCTTGCGAGTCTGAGACATCGGGCCAGAGTGTCCGCGCTAGG	7512
Db	8101	AAACTTTGGGTTACCGCCCTTGCGAGTCTGAGACATCGGGCCAGAGTGTCCGCGCTAGG	8160
QY	7513	CTACTGTCCAGGGGGGAGGGCTGCGACTCTGTGGCAAGTACCTCTTCAACTGGGCGAGTA	7572
Db	8161	CTACTGTCCAGGGGGGAGGGCTGCGACTCTGTGGCAAGTACCTCTTCAACTGGGCGAGTA	8220
QY	7573	AGGACCAAGCTCAAATCACTCCAAATCCCGGCTGCGTCCCAAGTTGGATTTATTCAGCTGG	7632
Db	8221	AGGACCAAGCTCAAATCACTCCAAATCCCGGCTGCGTCCCAAGTTGGATTTATTCAGCTGG	8280
QY	7633	TTTGGTTGCTGGTTACAGCGGGGGAGACATATATCACAGCTGTCTGTCGCCGACCCCGC	7692
Db	8281	TTTGGTTGCTGGTTACAGCGGGGGAGACATATATCACAGCTGTCTGTCGCCGACCCCGC	8340



1009 ACTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATA 1068  
1021 ACTGTGCGCGCTGGGTGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATA 1080  
1069 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCG 1128  
1081 TTGCTGAAGAGCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCG 1140  
1129 CTCCCGATTGCGAGCGGATCGCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAA 1188  
1141 CTCCCGATTGCGAGCGGATCGCTTCTATCGCCTTCTTGACGAGTCTTCTGAGTTTAAA 1200  
1189 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCT 1248  
1201 CAGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCT 1260  
1249 AAGCTTACTGGCGGAGCGCTTGGAAATAGGCGCGGTGCGTTGTGCTATATGTTATTT 1308  
1261 AAGCTTACTGGCGGAGCGCTTGGAAATAGGCGCGGTGCGTTGTGCTATATGTTATTT 1320  
1309 TCACCATATGTCGCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGCTTCTTG 1368  
1321 TCACCATATGTCGCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGCTTCTTG 1380  
1369 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTC 1428  
1381 ACGAGCATTCCTAGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTC 1440  
1429 GTGAAGGAAGCAGTTCCTCTGGAAGTCTTGTGAAGACAAACACCTCTGAGCGACCTT 1488  
1441 GTGAAGGAAGCAGTTCCTCTGGAAGTCTTGTGAAGACAAACACCTCTGAGCGACCTT 1500  
1489 TGCAAGCAGCGGAACCCCTCTGCGGACAGGTGCTCTGCGGCCAAAAAGCCACGTTA 1548  
1501 TGCAAGCAGCGGAACCCCTCTGCGGACAGGTGCTCTGCGGCCAAAAAGCCACGTTA 1560  
1549 TAAGATACCTGCAAGGGGCGACACCCAGTGCACCGTCTGAGTGGATAGTTGTG 1608  
1561 TAAGATACCTGCAAGGGGCGACACCCAGTGCACCGTCTGAGTGGATAGTTGTG 1620  
1609 GAAAGAGTCAAAATGGCTCTCTCAAGGATATCAACAAAGGGGTGAAGGATGCCAGAAG 1668  
1621 GAAAGAGTCAAAATGGCTCTCTCAAGGATATCAACAAAGGGGTGAAGGATGCCAGAAG 1680  
1669 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCAATGCTTTAATGTTTAG 1728  
1681 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCAATGCTTTAATGTTTAG 1740  
1729 TCAGAGTTAAAAACGTTAGGCCCCCGGAAACACCGGAGCGTGTTCCTTTGAAAAA 1788  
1741 TCAGAGTTAAAAACGTTAGGCCCCCGGAAACACCGGAGCGTGTTCCTTTGAAAAA 1800  
1789 CACGATAATACCATG----- 1803  
1801 CACGATAATACCATGACCGGAGATGGAGCATCTGTGCGGAGCGCGGTTTTCGTAGGT 1860  
1804 ----- 1803  
1861 CTGATACTCTTGACTTGTACCGCATATAAGTGTTCCTCGTAGGCTCATATGGTGG 1920  
1804 ----- 1803  
1921 TTACAATATTTTATCAGAGGCGGAGGACACTTGTGAAGTGTGGATCCCGCCCTCAAC 1980  
1804 ----- 1803  
1981 GTTCGGGGGGCGCGATGCGGTATCTCTCCTCAGCTGCGCGATCCACCGAGAGTAAATC 2040  
1804 ----- 1803  
2041 TTTTACCATCACCAAAATCTTGCTCGGCATACTCGGTCCACTCATGTGTCTCCAGGCTGGT 2100

1804 ----- 1803  
2101 ATAACAAAGTGGCGTACTTGTGCGGACACAGGGCTCATTCGTGATGATGCTGCTG 2160  
1804 ----- 1803  
2161 CGGAAGTTGCTGGGGTCAATTATGTCCAATGGCTCTCATGAAGTTGGCGGCACTGACA 2220  
1804 ----- 1803  
2221 GSTACGTACGTTTATGACCATCTCACCCACTGCGGACTGGGCCACGCGGCGCTACGA 2280  
1804 ----- 1803  
2281 GACCTTGGGTGCGAGTTGAGCCCGTGTCTTCTCTGATATGAGAGACCAAGTTTATCACC 2340  
1804 ----- 1803  
2341 TGGGGGACAGACACCGCGCGGTGTGGGACATCATCTTGGGCGCTGCCCTCTCGCCCGC 2400  
1804 ----- 1803  
2401 AGGGGAGGAGATACATCTGGGACCGGACAGACAGCCCTTGAAAGGAGGCTGCGCACTC 2460  
1804 ----- 1803  
2461 CTGCGCCTATTACGGCCTACTCCCAACAGACGCGAGGCTACTTGGCTGCTCATCACT 2520  
1861 AGCTCACAGCGCGGACAGGAAACAGGTCGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
2521 AGCTCACAGCGCGGACAGGAAACAGGTCGAGGGGAGGTCCAAAGTGTCTCCACCGCA 2580  
1921 ACACAATCTTCTGCGGACCTCGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGTC 1980  
2581 ACACAATCTTCTGCGGACCTCGCTCAATGGCGTGTGTTGGACTGTCTATCATGTGTC 2640  
1981 GGCTCAAGAGACCTTGGCGGCCCAAGGGGCCAATCACCCAAATGTACCAATGTGGAC 2040  
2641 GGCTCAAGAGACCTTGGCGGCCCAAGGGGCCAATCACCCAAATGTACCAATGTGGAC 2700  
2041 CAGGACCTCGCTGGTGGCAAGCGCCCGCGGCGGCTTCTTGACACCATGCACTGTC 2100  
2701 CAGGACCTCGCTGGTGGCAAGCGCCCGCGGCGGCTTCTTGACACCATGCACTGTC 2760  
2101 GGAGCTGCGGACCTTTACTTGTGTGTCAGAGGATGCGGATGCTCATTCGCTGCGCGCGG 2160  
2761 GGAGCTGCGGACCTTTACTTGTGTGTCAGAGGATGCGGATGCTCATTCGCTGCGCGCGG 2820  
2161 GGCGACAGGAGGAGGAGCTTCTCTCCCGAGCGCGTCTCTACTTGAAGGCTCTTTCG 2220  
2821 GGCGACAGGAGGAGGAGGAGCTTCTCTCCCGAGCGCGTCTCTACTTGAAGGCTCTTTCG 2880  
2221 GGCGGTCCTGCTCTGCGGCGAGCTGTGGGATCTTTCGGGCTGCGGCTGCTGTC 2280  
2881 GGCGGTCCTGCTCTGCGGCGAGCTGTGGGATCTTTCGGGCTGCGGCTGCTGTC 2340  
2281 ACCGAGGAGGTTGCGAAGCGGTGGAATTTGTATACCGCTCGAGTCTATGGAACCACTATG 2340  
2941 ACCGAGGAGGTTGCGAAGCGGTGGAATTTGTATACCGCTCGAGTCTATGGAACCACTATG 3000  
2341 CGGTCCCGGCTCTTACGAGCAACTCTGTCCTCCCGCGTACCGGACATTCAGGTTG 2400  
3001 CGGTCCCGGCTCTTACGAGCAACTCTGTCCTCCCGCGTACCGGAGCAATTCAGGTTG 3060  
2401 GCCCATCTACGCGCCCTACTGTAGCGCAAGAGCACTAAAGTGCCTGCGGCTGGTATGCA 2460  
3061 GCCCATCTACGCGCCCTACTGTAGCGCAAGAGCACTAAAGTGCCTGCGGCTGGTATGCA 3120  
2461 GCCCAAGGATATAAGGTGCTTGTCTGAACCGGTCGTCGCGGCCACCCCTAGGTTTCGGG 2520  
3121 GCCCAAGGATATAAGGTGCTTGTCTGAACCGGTCGTCGCGGCCACCCCTAGGTTTCGGG 3180  
2521 GCGTATATCTTAAGGCAATGATTCGACCCCTAAATCATCAGAACCGGGGTAAAGGACATC 2580





QY 4741 GTCCCTCTTCTCATGTCAACGTGGGTACAAAGGAGTCTGGGGGCGACGCGCATCATG 4800  
 Db 5401 GTCCCTCTTCTCATGTCAACGTGGGTACAAAGGAGTCTGGGGGCGACGCGCATCATG 5460  
 QY 4801 CAAACACCTGCCATGTGGAGACACAGATCACCGGACATGTGAAAAAGGTTCCATGAGG 4860  
 Db 5461 CAAACACCTGCCATGTGGAGACACAGATCACCGGACATGTGAAAAAGGTTCCATGAGG 5520  
 QY 4861 ATCGTGGGCTTAGGACCTGTAGTAACACGTGSCATGGAAATTCGCCATTAACCGGTAC 4920  
 Db 5521 ATCGTGGGCTTAGGACCTGTAGTAACACGTGSCATGGAAATTCGCCATTAACCGGTAC 5580  
 QY 4921 ACCACGGCCCTCGACGGCCCTCCCGGGCCCAATTAATCTAGGGCGCTGTGGCGGGTG 4980  
 Db 5581 ACCACGGCCCTCGACGGCCCTCCCGGGCCCAATTAATCTAGGGCGCTGTGGCGGGTG 5640  
 QY 4981 GCTGCTGAGGAGTACGTGGAGGTTACGGGGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
 Db 5641 GCTGCTGAGGAGTACGTGGAGGTTACGGGGGTGGGGATTTCCACTACGTGACGGGCATG 5700  
 QY 5041 ACCACTGACAAAGTAAAGTCCCGTGTAGGTTCCGGGCCCGGAAATTCCTCAAGAGTG 5100  
 Db 5701 ACCACTGACAAAGTAAAGTCCCGTGTAGGTTCCGGGCCCGGAAATTCCTCAAGAGTG 5760  
 QY 5101 GATGGGTGGGTGACACAGGTACGTCACAGCTGCAAGCGTCAAAACCCCTCTACGGGAGGTC 5160  
 Db 5761 GATGGGTGGGTGACACAGGTACGTCACAGCTGCAAGCGTCAAAACCCCTCTACGGGAGGTC 5820  
 QY 5161 ACATTCCTGCTGGGTCAATCAATACCTGGTGGGTACAGCTCCCATCCGAGCCCGAA 5220  
 Db 5821 ACATTCCTGCTGGGTCAATCAATACCTGGTGGGTACAGCTCCCATCCGAGCCCGAA 5880  
 QY 5221 CCGGACGTAGCTGCTCACTTCATGCTCAACGACCCCTCCACATTAACGGCGGAGAG 5280  
 Db 5881 CCGGACGTAGCTGCTCACTTCATGCTCAACGACCCCTCCACATTAACGGCGGAGAG 5940  
 QY 5281 GCTAAGCGTAGGCTGGCAGGGATCTCCCTCCCTCTTGGCCAGCTCATCAGTAGCCAG 5340  
 Db 5941 GCTAAGCGTAGGCTGGCAGGGATCTCCCTCCCTCTTGGCCAGCTCATCAGTAGCCAG 6000  
 QY 5341 CTGCTGCGCTTCCTTGAAGGCAACATGACCTACCTGCTACCTCCCGGACGCTGAC 5400  
 Db 6001 CTGCTGCGCTTCCTTGAAGGCAACATGACCTACCTGCTACCTCCCGGACGCTGAC 6060  
 QY 5401 CTATCGAGCCAACTCTCTGCGCGCAGGAGATGGCGGGAACATCAACCCGCTGGAG 5460  
 Db 6061 CTATCGAGCCAACTCTCTGCGCGCAGGAGATGGCGGGAACATCAACCCGCTGGAG 6120  
 QY 5461 TCAGAAATTAAGTAGTAATTTGGACTCTTCGAGCCGCTCCAGGGGAGGATGAG 5520  
 Db 6121 TCAGAAATTAAGTAGTAATTTGGACTCTTCGAGCCGCTCCAGGGGAGGATGAG 6180  
 QY 5521 AGGGAAGTATCCCTTCCGGCGAGATCTTCGAGAGTTCAGGAAATTCCTCCAGCGATG 5580  
 Db 6181 AGGGAAGTATCCCTTCCGGCGAGATCTTCGAGAGTTCAGGAAATTCCTCCAGCGATG 6240  
 QY 5581 CCATATGGGCAACCGCGGATTAACACCTCCACTGTTAGAGTCTCTGGAAAGGACCCGGAC 5640  
 Db 6241 CCATATGGGCAACCGCGGATTAACACCTCCACTGTTAGAGTCTCTGGAAAGGACCCGGAC 6300  
 QY 5641 TAGCTCCCTCCAGGTGACACGGGTGTCATTCGCGCTGCGAAGGCCCTCCGATACCA 5700  
 Db 6301 TAGCTCCCTCCAGGTGACACGGGTGTCATTCGCGCTGCGAAGGCCCTCCGATACCA 6360  
 QY 5701 CCTCCAGGAGGAGGAGCGGTGCTCTGTCAGAAATCTACCGTGTCTTCGCTTGGCG 5760  
 Db 6361 CCTCCAGGAGGAGGAGCGGTGCTCTGTCAGAAATCTACCGTGTCTTCGCTTGGCG 6420  
 QY 5761 GAGCTGCCACAAAGACCTTCGCGAGCTCCGAATCGTCGCGCTCGACAGCGGACCGCA 5820  
 Db 6421 GAGCTGCCACAAAGACCTTCGCGAGCTCCGAATCGTCGCGCTCGACAGCGGACCGCA 6480

QY 5821 AGGCTCTCTTGACAGAGCCCTCCGAGCGGCGACGCGGATCCGAGCTTGAGTCGTAC 5880  
 Db 6481 AGGCTCTCTTGACAGAGCCCTCCGAGCGGCGACGCGGATCCGAGCTTGAGTCGTAC 6540  
 QY 5881 TCTCCTCATGCCCTTGTGAGGGGAGCGGGGGATCCCGATCTCAGCGGACGGGTCTTGG 5940  
 Db 6541 TCTCCTCATGCCCTTGTGAGGGGAGCGGGGGATCCCGATCTCAGCGGACGGGTCTTGG 6600  
 QY 5941 TCTACCGTAAAGGAGGAGCTAGTAGAGAGCTGTGTGCTCGATGCTCTACATG 6000  
 Db 6601 TCTACCGTAAAGGAGGAGCTAGTAGAGAGCTGTGTGCTCGATGCTCTACATG 6660  
 QY 6001 ACAGGCGCCCTGATCAGCCATCGCTGCGGAGAAACAAAGCTGCCATCAATGACATG 6060  
 Db 6661 ACAGGCGCCCTGATCAGCCATCGCTGCGGAGAAACAAAGCTGCCATCAATGACATG 6720  
 QY 6061 AGCAACTCTTTGCTCGCTCACCACAACTTGGTCTATGCTACAACTCTCGACGCAAGC 6120  
 Db 6721 AGCAACTCTTTGCTCGCTCACCACAACTTGGTCTATGCTACAACTCTCGACGCAAGC 6780  
 QY 6121 CTGCGGACAGAGGATCACTTTTGACAGACTGCGAGTCTTGACAGCACTACCGGAC 6180  
 Db 6781 CTGCGGACAGAGGATCACTTTTGACAGACTGCGAGTCTTGACAGCACTACCGGAC 6840  
 QY 6181 GTGCTCAAGGAGATGAAGCGAGGCTCCACAGTTAAAGCTAAACTTCTATCCGTGGAG 6240  
 Db 6841 GTGCTCAAGGAGATGAAGCGAGGCTCCACAGTTAAAGCTAAACTTCTATCCGTGGAG 6900  
 QY 6241 GAAGCTGTAAAGCTGACGCGCCCACTTCGGCCAGATCTAAATTTGGCTATGGGCAAG 6300  
 Db 6901 GAAGCTGTAAAGCTGACGCGCCCACTTCGGCCAGATCTAAATTTGGCTATGGGCAAG 6960  
 QY 6301 GAGCTCCGAACTATCCAGCAAGCGGTTAAACCAATCCGCTCCGTGGAGAGACTTG 6360  
 Db 6961 GAGCTCCGAACTATCCAGCAAGCGGTTAAACCAATCCGCTCCGTGGAGAGACTTG 7020  
 QY 6361 CTGGAGACACTGAGACCAATTTGACACCACTCATCGGCAAAATGAGTTTCTGC 6420  
 Db 7021 CTGGAGACACTGAGACCAATTTGACACCACTCATCGGCAAAATGAGTTTCTGC 7080  
 QY 6421 GTCCAAACAGAGAGGGGCGCGCAAGCCAGCTCGGCTTATCGTATTCACAGATTTGGGG 6480  
 Db 7081 GTCCAAACAGAGAGGGGCGCGCAAGCCAGCTCGGCTTATCGTATTCACAGATTTGGGG 7140  
 QY 6481 GTTCTGTGTGCGAGAAATGCGCTTTACAGATGCTTCCACCTCCCTCAGGCGGTG 6540  
 Db 7141 GTTCTGTGTGCGAGAAATGCGCTTTACAGATGCTTCCACCTCCCTCAGGCGGTG 7200  
 QY 6541 ATGGGCTCTTACCGATTTCCAAATCTCTCTGGACAGGGGTGAGTCTCTGCTGAAT 6600  
 Db 7201 ATGGGCTCTTACCGATTTCCAAATCTCTCTGGACAGGGGTGAGTCTCTGCTGAAT 7260  
 QY 6601 GCTGGAAGCGAGAAATGCGCTTATGGGCTTCGCATATGACACCGGCTGTTTGTACTCA 6660  
 Db 7261 GCTGGAAGCGAGAAATGCGCTTATGGGCTTCGCATATGACACCGGCTGTTTGTACTCA 7320  
 QY 6661 AGGTCACCTAGAAATGACATCCGTGTTGAGAGTCAATCTACAAATGTTGTGCTGGCC 6720  
 Db 7321 AGGTCACCTAGAAATGACATCCGTGTTGAGAGTCAATCTACAAATGTTGTGCTGGCC 7380  
 QY 6721 CCGGAGCCAGACAGCCATAGGTCGCTCAGAGCGGCTTTACATCGGGGCCCCCTG 6780  
 Db 7381 CCGGAGCCAGACAGCCATAGGTCGCTCAGAGCGGCTTTACATCGGGGCCCCCTG 7440  
 QY 6781 ACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
 Db 7441 ACTAATTTAAAGGCGAGAACTGCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 7500  
 QY 6841 ACCAGTCGCGTAATACCTCAGATGTTACTTTGAAGGCGGCTGCGGCTGTGAGCTGGG 6900  
 Db 7501 ACCAGTCGCGTAATACCTCAGATGTTACTTTGAAGGCGGCTGCGGCTGTGAGCTGGG 7560  
 QY 6901 AAGCTCCAGGACTGACAGATGCTGATCGGAGACGACCTTGTGCTGTAAGC 6960

```
Db 7561 AAGCTCCAGGACTGCAACGATGCTCGTATGCGGAGACGACCTTGTGCTTATCTGTGAAGC 7620
Qy 6961 GCGGGAGCCCAAGAGGAGGAGCGAGCTTA CGGGCTTTCACGGAGGCTATGACTAGATAC 7020
Db 7621 GCGGGAGCCCAAGAGGAGGAGCGAGCTTA CGGGCTTTCACGGAGGCTATGACTAGATAC 7680
Qy 7021 TCTGCCCCCTGGGGAGCCCGCCAAACAGAAATACGACTTGGAGTTGATATACATCATGC 7080
Db 7681 TCTGCCCCCTGGGGAGCCCGCCAAACAGAAATACGACTTGGAGTTGATATACATCATGC 7740
Qy 7081 TCTTCCAATGTGTAGTCGCGACGATGATCTGGCAAAAGGCTGACTATCTCAACCGT 7140
Db 7741 TCTTCCAATGTGTAGTCGCGACGATGATCTGGCAAAAGGCTGACTATCTCAACCGT 7800
Qy 7141 GACCCACACACCCCTTGGCGGCTGGTGGGAGACGCTAGACACATCCAGTCAAT 7200
Db 7801 GACCCACACACCCCTTGGCGGCTGGTGGGAGACGCTAGACACATCCAGTCAAT 7860
Qy 7201 TCTGGCTAGGCAACATCATGATGCGCCACCTTGTGGCAAGGATGATCTGTATG 7260
Db 7861 TCTGGCTAGGCAACATCATGATGCGCCACCTTGTGGCAAGGATGATCTGTATG 7920
Qy 7261 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGA AAAAGCCCTAGATTGTCA 7320
Db 7921 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGA AAAAGCCCTAGATTGTCA 7980
Qy 7321 ATCTACGGGGCTGTATTCTCATGAGGACCTTGTAGCTACCTACATCATCAAGACTC 7380
Db 7981 ATCTACGGGGCTGTATTCTCATGAGGACCTTGTAGCTACCTACATCATCAAGACTC 8040
Qy 7381 CATGGCTTAGGCAATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCT 7440
Db 8041 CATGGCTTAGGCAATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCT 8100
Qy 7441 TCATGCTCAGGAACCTTGGGCTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAGT 7500
Db 8101 TCATGCTCAGGAACCTTGGGCTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAGT 8160
Qy 7501 GTCCGCGCTAGGCTACTGTCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560
Db 8161 GTCCGCGCTAGGCTACTGTCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 8220
Qy 7561 AACTGGGAGTAAGACCAAGCTCAAACTCACTCCAACTCCGGCTGCGTCCAGTTGGAT 7620
Db 8221 AACTGGGAGTAAGACCAAGCTCAAACTCACTCCAACTCCGGCTGCGTCCAGTTGGAT 8280
Qy 7621 TTATCCAGCTGGTTCTGTTCTGTTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7680
Db 8281 TTATCCAGCTGGTTCTGTTTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 8340
Qy 7681 GCCGACCCCGCTGTTTCACTGTTGCTTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 7740
Db 8341 GCCGACCCCGCTGTTTCACTGTTGCTTACTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 8400
Qy 7741 CTACTCCCCAACCCGATGAACGGGACCTAAACACTCCAGGCAATAGGCCATCTCTGTTT 7800
Db 8401 CTACTCCCCAACCCGATGAACGGGAGCTAAACACTCCAGGCAATAGGCCATCTCTGTTT 8460
Qy 7801 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7860
Db 8461 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 8520
Qy 7861 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 7920
Db 8521 TTTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 8580
Qy 7921 TAGCTGTGAAGGTCGGTGAGCGGCTTGA CTGAGAGAGTGTGATAGTGGCTCTCTGC 7980
Db 8581 TAGCTGTGAAGGTCGGTGAGCGGCTTGA CTGAGAGAGTGTGATAGTGGCTCTCTGC 8640
Qy 7981 AGATCAAGT 7989
|||||
```

```
Db 8641 AGATCAAGT 8649
RESULT 8
US-10-029-907-1
; Sequence 1, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: HEPATITIS C VIRUS
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 8639
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1803) ... (8408)
US-10-029-907-1
Query Match 91.0%; Score 7272.2; DB 4; Length 8639;
Best Local Similarity 92.2%; Pred. No. 0;
Matches 7975; Conservative 0; Mismatches 3; Indels 671; Gaps 3;
Qy 1 GCCAGCCCCGATTCGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 60
Db 2 GCCAGCCCCGATTCGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAATCTACTG 61
Qy 61 TCTTCACGAGAAAGCGTCTAGCCATGCGTGTAGTATGAGTGTGTCGAGCCTCCAGGAC 120
Db 62 TCTTCACGAGAAAGCGTCTAGCCATGCGTGTAGTATGAGTGTGTCGAGCCTCCAGGAC 121
Qy 121 CCCCCCTCCCGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATTCGCG 180
Db 122 CCCCCCTCCCGGAGAGCCATAGTGGTCTGGGAAACCGGTGAGTACACCGGAATTCGCG 181
Qy 181 GAGGACCGGGTCTTTTCTTTGGATCAACCGCTCAATGCTCGAGATTTGGGCGTCCGCC 240
Db 182 GAGGACCGGGTCTTTTCTTTGGATCAACCGCTCAATGCTCGAGATTTGGGCGTCCGCC 241
Qy 241 GCGAGACTGCTAGCCGAGTGTGTTGGTTCGCGAAAGGCTTGTGTTACTGCTGATAGG 300
Db 242 GCGAGACTGCTAGCCGAGTGTGTTGGTTCGCGAAAGGCTTGTGTTACTGCTGATAGG 301
Qy 301 GTGCTTGGAGTGCCTCCGGAGGTTCTCGTAGCCGTGACCATGAGCAGCATCTTAAC 360
Db 302 GTGCTTGGAGTGCCTCCGGAGGTTCTCGTAGCCGTGACCATGAGCAGCATCTTAAC 361
Qy 361 CTCAGAAAGAAACCAAGGGCGCGCATGATTGAACAAGATGGATTGACGAGGTTCTC 420
Db 362 CTCAGAAAGAAACCAAGGGCGCGCATGATTGAACAAGATGGATTGACGAGGTTCTC 421
Qy 421 CGGCGCTTGGGTGGAGAGGCTATTCGGCTATGATGGGCAACACAGACAATTCGGCTGT 480
Db 422 CGGCGCTTGGGTGGAGAGGCTATTCGGCTATGATGGGCAACACAGACAATTCGGCTGT 481
Qy 481 CTGATGCGCGCTGTTTCCGGCTGTGAGCGGCGGCGCGGCTTCTTTTGTCAAGACG 540
Db 482 CTGATGCGCGCTGTTTCCGGCTGTGAGCGGCGGCGGCTTCTTTTGTCAAGACG 541
Qy 541 ACCTCTCGGTCGCCCTGAATCAACTGACGAGGACGAGCGCGGCTATCGTGGCTGGCA 600
Db 542 ACCTCTCGGTCGCCCTGAATCAACTGACGAGGACGAGCGCGGCTATCGTGGCTGGCA 601
Qy 601 CGACGGGCTTCTTTCGCGAGCTGTGTCGAGTTGTCACTGAAGCGGGAAGGAGTGGC 660
|||||
```

Db 602 CGAGGGGCGTTCTTGGCGAGCTGTCTCGACGTCTCACTGAAGCGGGAAGGACTGGC 661  
QY 661 TGGTATTGGGCGAAGTGGCGGGGCAAGATCTCTGTCACTCTCACTTGTCTCTGGCGGAGA 720  
Db 662 TGGTATTGGGCGAAGTGGCGGGGCAAGATCTCTGTCACTCTCACTTGTCTCTGGCGGAGA 721  
QY 721 AAGTATCCATCATGGCTGATGCAATCGGCGGCTGCATACGCTTGATCGGCTACCTGGC 780  
Db 722 AAGTATCCATCATGGCTGATGCAATCGGCGGCTGCATACGCTTGATCGGCTACCTGGC 781  
QY 781 CATTCGACCACCAAGCGAATCGCATCGAGCGAGCAGTACTCGGATGGAAGCCGGTC 840  
Db 782 CATTCGACCACCAAGCGAATCGCATCGAGCGAGCAGTACTCGGATGGAAGCCGGTC 841  
QY 841 TTGTGATCAGGATGATCTGGAGGAAGAGATCAGGGGCTCGGCGCAGCGAATCTTTGG 900  
Db 842 TTGTGATCAGGATGATCTGGAGGAAGAGATCAGGGGCTCGGCGCAGCGAATCTTTGG 901  
QY 901 CCAGGCTCAAGGCGGCATGCCGAGCGGAGGATCTCGTGTGACCCATGGCGATGCCCT 960  
Db 902 CCAGGCTCAAGGCGGCATGCCGAGCGGAGGATCTCGTGTGACCCATGGCGATGCCCT 961  
QY 961 GCTTGC CGAATATCATGCTGGAAATGGCGCTTTTCTGGAATCATCGACTGTGGCGGC 1020  
Db 962 GCTTGC CGAATATCATGCTGGAAATGGCGCTTTTCTGGAATCATCGACTGTGGCGGC 1021  
QY 1021 TGGGCTGGCGGACCGCTATCAGGATAGAGCTGGCTTGGCTTACCGCTGATATTTGTAAGGC 1080  
Db 1022 TGGGCTGGCGGACCGCTATCAGGATAGAGCTGGCTTGGCTTACCGCTGATATTTGTAAGGC 1081  
QY 1081 TTGGCGGCGAATGGGCTGACCGCTTCTGTGCTTTACGGTATCGCGCTCCCGATTCGC 1140  
Db 1082 TTGGCGGCGAATGGGCTGACCGCTTCTGTGCTTTACGGTATCGCGCTCCCGATTCGC 1141  
QY 1141 AGCGCATCGCTTCTATCGCTTTCTTGACGAGTCTTCTGAGTT-----TAAA 1188  
Db 1142 AGCGCATCGCTTCTATCGCTTTCTTGACGAGTCTTCTGAGTTTCGGCGCCAGATGTTAA 1201  
QY 1189 CAGACCAACAGGTTTCCCTTAGCGGATCAATTCGCGCCCTCTCCCTTCCCGCCCGCTT 1248  
Db 1202 CAGACCAACAGGTTTCCCTTAGCGGATCAATTCG-----CCCCCCCCCT 1250  
QY 1249 AAGCTTACTGGCGGAGCGCTTGGATAGCGCGGTGTGCTTGTCTATATGTTATTT 1308  
Db 1251 AAGCTTACTGGCGGAGCGCTTGGATAGCGCGGTGTGCTTGTCTATATGTTATTT 1310  
QY 1309 TCCACCATATTCGCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGTCTTCTTG 1368  
Db 1311 TCCACCATATTCGCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCCCTGTCTTCTTG 1370  
QY 1369 ACAGCATTCCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTC 1428  
Db 1371 ACAGCATTCCTAGGGGTCTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTC 1430  
QY 1429 GTGAAGGAAGCTTCTCTGGAAGCTTCTTGAAGCAACAACGCTGTAGCGACCTT 1488  
Db 1431 GTGAAGGAAGCTTCTCTGGAAGCTTCTTGAAGCAACAACGCTGTAGCGACCTT 1490  
QY 1489 TGAGGAGCGGGAACCCCGCTGCGGACAGGTGCTCTGCGGCCAAAAGCCACCTGTA 1548  
Db 1491 TGAGGAGCGGGAACCCCGCTGCGGACAGGTGCTCTGCGGCCAAAAGCCACCTGTA 1550  
QY 1549 TAAGATACACTGGAAGGCGCAACCCAGTGCACGCTTGTGAGTTGGAATGTTG 1608  
Db 1551 TAAGATACACTGGAAGGCGCAACCCAGTGCACGCTTGTGAGTTGGAATGTTG 1610  
QY 1609 GAAGAGTCAATGGCTCTCTCAAGGATTTCAAGAGGCTGAGGATGCCAGAG 1668  
Db 1611 GAAGAGTCAATGGCTCTCTCAAGGATTTCAAGAGGCTGAGGATGCCAGAG 1670  
QY 1669 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTGTTAG 1728  
Db 1671 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTGTTAG 1730

QY 1729 TCGAGGTTAAAAACGTTAGGCCCCCGAACCCACGGGACGTTGTTTCTTTGAAAAA 1788  
Db 1731 TCGAGGTTAAAAACGTTAGGCCCCCGAACCCACGGGACGTTGTTTCTTTGAAAAA 1790  
QY 1789 CAGGATAATACCATG----- 1803  
Db 1791 CAGGATAATACCATGACCGGGAGATGGCAGCATCTGTCGGAGCGGGTTTCGTAGGT 1850  
QY 1804 ----- 1803  
Db 1851 CTGATACTCTTGACCTTGTACCGCACTATAAGCTTCTCTCGCTAGGCTCATATGTTGG 1910  
QY 1804 ----- 1803  
Db 1911 TTACAATATTTTATCACCAGGCGGAGGCACACTTTGCAAGTGTGGATCCCGCCCTCAAC 1970  
QY 1804 ----- 1803  
Db 1971 GTTCGGGGGGCGGATGCGGTCTCTCTCACGTGCGGATCCACCAGAGCTAATC 2030  
QY 1804 ----- 1803  
Db 2031 TTTACCATCACAAAATCTTGTCTGCCATACTCGTCCACTCATGTGTCTCCAGGCTGT 2090  
QY 1804 ----- 1803  
Db 2091 ATAACAAAAGTCCGTPACTTGTGCGCGCACACGGGCTCATTCGTGATGATGCTGGTG 2150  
QY 1804 ----- 1803  
Db 2151 CGGAAGTTGCTGGGGTCAATTATGTCCAAATGGTCTCATGAAGTTGGCGCAGCTGACA 2210  
QY 1804 ----- 1803  
Db 2211 GGTACGTACGTTTATGACCATCTCACCCACTCTGCGGACTGCGGCCACCGCGCCCTACGA 2270  
QY 1804 ----- 1803  
Db 2271 GACCTTTCGGTGGCAGTTGAGCCCGTCTTCTCTGATATGAGAGCAAGGTTATCACC 2330  
QY 1804 ----- 1803  
Db 2331 TGGGGGGGAGACACCGCGCGGTGTGGGGACATCATCTTGGGCTGCGCGCTCTCGGCCCGC 2390  
QY 1804 ----- 1803  
Db 2391 AGGGGAGGGAGATACATCTGGGACCGGCAGACAGCCTTGAAGGGCAGGGGTGGCGACTC 2450  
QY 1804 ----GCGCCTATTACGGCTTACTCCGAAACAGACGCGAGGCTTACTTGGCTGCATCATCACT 1860  
Db 2451 CTGCGCCCTATTACGGCTTACTCCGAAACAGACGCGAGGCTTACTTGGCTGCATCATCACT 2510  
QY 1861 AGCCTCACAGGCGGACAGGAACACAGGTCTGAGGGGAGGTCCAAAGTGTCTCCACGCA 1920  
Db 2511 AGCCTCACAGGCGGACAGGAACACAGGTCTGAGGGGAGGTCCAAAGTGTCTCCACGCA 2570  
QY 1921 ACACAATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTGACTGTCTTATCATGTTGCC 1980  
Db 2571 ACACAATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTGACTGTCTTATCATGTTGCC 2630  
QY 1981 GGTCAAAGACCTTTCGCGCCCAAGGGCCCAATCACCAAAATGTACCAATGTGGAC 2040  
Db 2631 GGTCAAAGACCTTTCGCGCCCAAGGGCCCAATCACCAAAATGTACCAATGTGGAC 2690  
QY 2041 CAGGACTCTGTGGCTGGCAAGCGCCCGCGGCGCTTCTTGTACACCATGCACTGC 2100  
Db 2691 CAGGACTCTGTGGCTGGCAAGCGCCCGCGGCGCTTCTTGTACCAATGCACTGC 2750  
QY 2101 GGCAGCTCGGACCTTTACTTGGTTCAGGAGCATGCGGATGTCTATCCGTTGCGCGCGG 2160  
Db 2751 GGCAGCTCGGACCTTTACTTGGTTCAGGAGCATGCGGATGTCTATCCGTTGCGCGCGG 2810

2161 GGCGACAGAGGGGAGGCTACTCTCCCGCCAGCGCCGCTCTCTACTTGAAGGCTCTTCG 2220  
2811 GGCGACAGAGGGGAGGCTACTCTCCCGCCAGCGCCGCTCTCTACTTGAAGGCTCTTCG 2870  
2221 GGCGGTCCACTGCTCTGCTGCTCGGCGTCCGCGGACGCTGTGGGCACTTTTCGGGCTGCGGTGTC 2280  
2871 GGCGGTCCACTGCTCTGCTGCTCGGCGTCCGCGGACGCTGTGGGCACTTTTCGGGCTGCGGTGTC 2930  
2281 ACCGAGGGGTGCGAAGGGCGGTGACCTTTGTACCCGTCGAGTCTATGGAAACCACTATG 2340  
2931 ACCGAGGGGTGCGAAGGGCGGTGACCTTTGTACCCGTCGAGTCTATGGAAACCACTATG 2990  
2341 CGGTCCCGGCTTTCACGAGCAAACTCGTCCCGTCCGCGCGTACCGCAGACATTCAGAGTG 2400  
2991 CGGTCCCGGCTTTCAGCGACAACTCGTCCCGTCCGCGCGTACCGCAGACATTCAGAGTG 3050  
2401 GCCCATCTACAGCGCCCTACTGTTAGGGCAAGAGCACTAAGGTGCGCGGCTGCGTATGCA 2460  
3051 GCCCATCTACAGCGCCCTACTGTTAGGGCAAGAGCACTAAGGTGCGCGGCTGCGTATGCA 3110  
2461 GCCCAAGGTATAGGTGCTTCTGCTGAACCGCTCGCTCGCGCGCACCCCTAGGTTTCGGG 2520  
3111 GCCCAAGGTATAGGTGCTTCTGCTGAACCGCTCGCTCGCGCGCACCCCTAGGTTTCGGG 3170  
2521 GCGTATATGCTAAGGCACATGGTATCGACCCCTAAACATCAGAACCGGGGTAAAGCACTC 2580  
3171 GCGTATATGCTAAGGCACATGGTATCGACCCCTAAACATCAGAACCGGGGTAAAGCACTC 3230  
2581 ACCAGGGTGCCCGCATCAGTACTCCAGCTATGCGCAAGTTCCTTTGCGGACGCGGTGTC 2640  
3231 ACCAGGGTGCCCGCATCAGTACTCCAGCTATGCGCAAGTTCCTTTGCGGACGCGGTGTC 3290  
2641 TCTGGGCGCGCTATGACATCATATATGATGATGTCCTCACTCACTGACTCGACCACT 2700  
3291 TCTGGGCGCGCTATGACATCATATATGATGATGTCCTCACTCACTGACTCGACCACT 3350  
2701 ATCTGGGCATCGGCACAGTCTCTGCAAGCAAGCGAGAGCGGTGGAGCGGACTCTCTGCTG 2760  
3351 ATCTGGGCATCGGCACAGTCTCTGCAAGCAAGCGAGAGCGGTGGAGCGGACTCTCTGCTG 3410  
2761 CTGCGACCGCTACGCTCTCGGGATCGGTACCGTGCCACATTCGCAATTCGAGAGGTG 2820  
3411 CTGCGACCGCTACGCTCTCGGGATCGGTACCGTGCCACATTCGCAATTCGAGAGGTG 3470  
2821 GCTCTGTGACGACTGAGAAATCCCTTTTATGCAAGCAAGCCATCCCATCGAGACCACTC 2880  
3471 GCTCTGTGACGACTGAGAAATCCCTTTTATGCAAGCAAGCCATCCCATCGAGACCACTC 3530  
2881 AAGGGGGGAGGCACCTCATTTTCTGCCATTCCAGAAAGAAATGTGATGAGCTCGCGCG 2940  
3531 AAGGGGGGAGGCACCTCATTTTCTGCCATTCCAGAAAGAAATGTGATGAGCTCGCGCG 3590  
2941 AAGCTGTGCGGCTCGGACTCAATCTGTAGCATATTCAGGGGCTTGTATTCGCTC 3000  
3591 AAGCTGTGCGGCTCGGACTCAATCTGTAGCATATTCAGGGGCTTGTATTCGCTC 3650  
3001 ATACCAATAGCGGAGAGCTCAATCTGTAGCAACGAGCGCTCTAATCAGCGGCTTTACC 3060  
3651 ATACCAATAGCGGAGAGCTCAATCTGTAGCAACGAGCGCTCTAATCAGCGGCTTTACC 3710  
3061 GCGGATTTGCACTGAGTCACTGCAATCATGTGTACCCAGACAGTTCAGCTCAGC 3120  
3711 GCGGATTTGCACTGAGTCACTGCAATCATGTGTACCCAGACAGTTCAGCTCAGC 3770  
3121 CTGACCCGAGCTTACCATTTAGAGCAGCGCTGCGCAGAGCGCGGTGTACGCTCG 3180  
3771 CTGACCCGAGCTTACCATTTAGAGCAGCGCTGCGCAGAGCGCGGTGTACGCTCG 3830  
3181 CAGCGCGAGGAGGACTGTTAGGGGAGGATGGGCAATTTACAGTTTGTGACTCCAGGA 3240  
3831 CAGCGCGAGGAGGACTGTTAGGGGAGGATGGGCAATTTACAGTTTGTGACTCCAGGA 3890  
3241 GAAAGGCGCTCGGGCATGTTTCGATTCCTCGGTTCTGTGCGAGTGCTATGACGCGGGCTGT 3300

3891 GAAAGGCGCTCGGGCATGTTTCGATTCCTCGGTTCTGTGCGAGTGTCTATGACGCGGCTGT 3950  
3301 GCTTGGTACGAGCTCACCGCCGCGAGACCTCAGTTAGTTGCGGGCTTACTAAACACA 3360  
3951 GCTTGGTACGAGCTCACCGCCGCGAGACCTCAGTTAGTTGCGGGCTTACTAAACACA 4010  
3361 CCAGGTTGCGGCTGCGCAGGACCATCTGAGAGTCTGGGAGAGGCTTTTACAGGCTC 3420  
4011 CCAGGTTGCGGCTGCGCAGGACCATCTGAGAGTCTGGGAGAGGCTTTTACAGGCTC 4070  
3421 ACCCACATAGAGCGCCCATTTCTTGTCCAGACTAAGCAGGCGAGGAGCAAACTTCCCTTAC 3480  
4071 ACCCACATAGAGCGCCCATTTCTTGTCTGCGAGACTAAGCAGGCGAGGAGCAAACTTCCCTTAC 4130  
3481 CTGTAGCATACGAGCTACGCTGCGCAGGGCTCAGGCTCAGCTCCATCTCATCTGTGGAC 3540  
4131 CTGTAGCATACGAGCTACGCTGCGCAGGGCTCAGGCTCAGCTCCATCTCATCTGTGGAC 4190  
3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGCACGGGCAACGCCCTGTGTG 3600  
4191 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGCACGGGCAACGCCCTGTGTG 4250  
3601 TATAGGCTGGAGCGGTTCAAAACGAGTTTACTACACACACCCCATAAACCAATATCATC 3660  
4251 TATAGGCTGGAGCGGTTCAAAACGAGTTTACTACACACACCCCATAAACCAATATCATC 4310  
3661 ATGGCATCATGTCGGCTGACCTGGAGGTCTCAGAGCACTGGGTGCTGTAGCGGA 3720  
4311 ATGGCATCATGTCGGCTGACCTGGAGGTCTCAGAGCACTGGGTGCTGTAGCGGA 4370  
3721 GTCTAGCAGCTCTGGCGCGTATTGCTGCTGCAACAGGAGCGGTGCTCATTTGTGGGAGG 3780  
4371 GTCTAGCAGCTCTGGCGCGTATTGCTGCTGCAACAGGAGCGGTGCTCATTTGTGGGAGG 4430  
3781 ATCATCTTTGTCGGAAAGCGCGCCATCATTTCCCGACAGGGAAGTCTTTACCGGAGTTC 3840  
4431 ATCATCTTTGTCGGAAAGCGCGCCATCATTTCCCGACAGGGAAGTCTTTACCGGAGTTC 4490  
3841 GATGAGTGAAGAGTGGCGCTCAGCTCTCTTACATCGAACAGGGAATGAGCTCGCC 3900  
4491 GATGAGTGAAGAGTGGCGCTCAGCTCTCTTACATCGAACAGGGAATGAGCTCGCC 4550  
3901 GAACAATTCAAACAGAGGCAATCGGTTGCTGAAACAGCCACCAAGCAGCGAGGCT 3960  
4551 GAACAATTCAAACAGAGGCAATCGGTTGCTGAAACAGCCACCAAGCAGCGAGGCT 4610  
3961 GCTGCTCCGCTGGTGAATCCAAAGTGGCGGACCTCTGAAAGCTTTCTGGGCGAAGCATATG 4020  
4611 GCTGCTCCGCTGGTGAATCCAAAGTGGCGGACCTCTGAAAGCTTTCTGGGCGAAGCATATG 4670  
4021 TGGAAATTCATCAGGGGATACATATTTAGAGGCTTGTGCACTCTGCTGGGCAACCC 4080  
4671 TGGAAATTCATCAGGGGATACATATTTAGAGGCTTGTGCACTCTGCTGGGCAACCC 4730  
4081 GCGATAGCATCACTGATGGCATTCACGCTCTATCAGCGCGCTCACCACCCACAT 4140  
4731 GCGATAGCATCACTGATGGCATTCACGCTCTATCAGCGCGCTCACCACCCACAT 4790  
4141 ACCCTCTGTTTAAACATCTCGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4200  
4791 ACCCTCTGTTTAAACATCTCGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCT 4850  
4201 GCTTCTGCTTTCGTAGGCGCGGATCGCTGGAGCGGTGTGCGAGCATAGGCTTGGG 4260  
4851 GCTTCTGCTTTCGTAGGCGCGGATCGCTGGAGCGGTGTGCGAGCATAGGCTTGGG 4910  
4261 AAGTGTCTGTGGATATTTTGGCAGTTATGGAGCAGGGGTGGCAGCGCTCTGGCC 4320  
4911 AAGTGTCTGTGGATATTTTGGCAGTTATGGAGCAGGGGTGGCAGCGCTCTGGCC 4970  
4321 TTTAAGGTCTAGGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTTACTCTCTGCT 4380

Db 4971 TTTAAGGTCAATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 5030  
QY 4381 ATCTCTCCCTTGGCGCCCTAGTCTGTCGGGTCGTGTGCGCAGGATACCTGCGTGGCAC 4440  
Db 5031 ATCTCTCCCTTGGCGCCCTAGTCTGTCGGGTCGTGTGCGCAGGATACCTGCGTGGCAC 5090  
QY 4441 GTGGGCCAGGGAGGGGCTGTGCAGTGTGAACCGGTCGTGTAGGTTGCGTTCGCGG 4500  
Db 5091 GTGGGCCAGGGAGGGGCTGTGCAGTGTGAACCGGTCGTGTAGGTTGCGTTCGCGG 5150  
QY 4501 GGTAAACCACTCTCCCGCACGCACTATGTGCTGTGAGAGCGAGCGTGTGAGTGTGCACT 4560  
Db 5151 GGTAAACCACTCTCCCGCACGCACTATGTGCTGTGAGAGCGAGCGTGTGAGTGTGCACT 5210  
QY 4561 CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGTGAAGAGGCTTCAACGTGTGAATCAAC 4620  
Db 5211 CAGATCTCTCTAGTCTTACCATCACTCAGCTGTGTGAAGAGGCTTCAACGTGTGAATCAAC 5270  
QY 4621 GAGGACTGTCTCCACGCCATCTCCGGCTGTGGCTAAGAGATGTTTGGGATTGGATATGC 4680  
Db 5271 GAGGACTGTCTCCACGCCATCTCCGGCTGTGGCTAAGAGATGTTTGGGATTGGATATGC 5330  
QY 4681 ACGGTGTGACTGATTCAAGACCTGCTCCAGTCCAGTCCAAAGCTCTGCGCGGATTCGCGGGA 4740  
Db 5331 ACGGTGTGACTGATTCAAGACCTGCTCCAGTCCAGTCCAAAGCTCTGCGCGGATTCGCGGGA 5390  
QY 4741 GTCCCTCTTCTCATGTCAACGTGGTCAAGGGAGTCTGGCGGGGCGACGCGATCATG 4800  
Db 5391 GTCCCTCTTCTCATGTCAACGTGGTCAAGGGAGTCTGGCGGGGCGACGCGATCATG 5450  
QY 4801 CAAACCACTGCGCATGTGGAGCACAGATCACCGGACATGTGAAACCGTTCCTATGAGG 4860  
Db 5451 CAAACCACTGCGCATGTGGAGCACAGATCACCGGACATGTGAAACCGTTCCTATGAGG 5510  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGCATGGAACATTCGCCATTAACGGGTAC 4920  
Db 5511 ATCGTGGGCTTAGGACCTGTAGTAAACAGTGCATGGAACATTCGCCATTAACGGGTAC 5570  
QY 4921 ACCACGGGCTTGCAGCGCTCCCGGGCCAAATTTATCTAGGGGCTGTGGCGGGTG 4980  
Db 5571 ACCACGGGCTTGCAGCGCTCCCGGGCCAAATTTATCTAGGGGCTGTGGCGGGTG 5040  
QY 4981 GTGTCTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
Db 5631 GCTGCTGAGGAGTACGTGGAGTTACGCGGTGGGGATTTCCACTACGTGACGGGCATG 5090  
QY 5041 ACCACTGACAAAGTAAAGTCCCGTGTCAAGTTCCCGCCCCCGAAATTTCTTCAAGAGTG 5100  
Db 5691 ACCACTGACAAAGTAAAGTCCCGTGTCAAGTTCCCGCCCCCGAAATTTCTTCAAGAGTG 5150  
QY 5101 GATGGGTCGGTGTGACAGGTACGTCCAGCGTGCAGACCCCTCTACGGGAGGAGTTC 5160  
Db 5751 GATGGGTCGGTGTGACAGGTACGTCCAGCGTGCAGACCCCTCTACGGGAGGAGTTC 5180  
QY 5161 ACATTCCTGCTCGGCTCAATCAATACCTGTTGGGTTCACAGCTCCCATGCGAGCCGGA 5220  
Db 5811 ACATTCCTGCTCGGCTCAATCAATACCTGTTGGGTTCACAGCTCCCATGCGAGCCGGA 5270  
QY 5221 CCGGAGTGTGAGTGTCTCACTTCCATGCTCACGACCCCTCCACATTAAGCGGAGAGC 5280  
Db 5871 CCGGAGTGTGAGTGTCTCACTTCCATGCTCACGACCCCTCCACATTAAGCGGAGAGC 5290  
QY 5281 GCTAAGGTTAGGCTGGCAGGGATCTCCCCCTCTTGGCAGCTCATCAGGTAGCCAG 5340  
Db 5931 GCTAAGGTTAGGCTGGCAGGGATCTCCCCCTCTTGGCAGCTCATCAGGTAGCCAG 5390  
QY 5341 CTGCTGTGCGCTTCTTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCTGAC 5400  
Db 5991 CTGCTGTGCGCTTCTTTGAAGGCAACATGCACTACCGTCACTGCTCCCGGAGCTGAC 6050  
QY 5401 CTCATGAGGCCAACCTCTGTGGCGGAGAGATGGGGGGAACATCAACCCGCTGGAG 5460  
Db 6051 CTCATGAGGCCAACCTCTGTGGCGGAGAGATGGGGGGAACATCAACCCGCTGGAG 6110

QY 5461 TCAGAAATAAGTAGTAAATTTTGGACTCTTTTCGAGCCGCTCCAGCGGAGGAGATGAG 5520  
Db 6111 TCAGAAATAAGTAGTAAATTTTGGACTCTTTTCGAGCCGCTCCAGCGGAGGAGATGAG 6170  
QY 5521 AGGGAATGATTCCTTCCGGCGGAGATCTTCGCGAGGTTCAGGAAATTTCCCTCCAGAGCATG 5580  
Db 6171 AGGGAATGATTCCTTCCGGCGGAGATCTTCGCGAGGTTCAGGAAATTTCCCTCCAGAGCATG 6230  
QY 5581 CCCATATGGGACCGCCCGGATTACAACCTCTCCACTGTTAGAGTCTTGGAAGGACCCGAC 5640  
Db 6231 CCCATATGGGACCGCCCGGATTACAACCTCTCCACTGTTAGAGTCTTGGAAGGACCCGAC 6290  
QY 5641 TAGCTCTCTCCAGTGTGTAACAGGCTGTCAATTCGCGCTGTCGAAAGGCCCTCCGATACCA 5700  
Db 6291 TAGCTCTCTCCAGTGTGTAACAGGCTGTCAATTCGCGCTGTCGAAAGGCCCTCCGATACCA 6350  
QY 5701 CCTCCACGGAGGAGGAGGTTGCTGTCTAGAAATCTACCGTGTCTTCTGCTGCTGGCG 5760  
Db 6351 CCTCCACGGAGGAGGAGGTTGCTGTCTAGAAATCTACCGTGTCTTCTGCTGCTGGCG 6410  
QY 5761 GAGCTCGCCACAAAGACCTTTCGGCAGCTCCGAATCGTCCGCCGTTCGACAGCGGACGCGCA 5820  
Db 6411 GAGCTCGCCACAAAGACCTTTCGGCAGCTCCGAATCGTCCGCCGTTCGACAGCGGACGCGCA 6470  
QY 5821 ACGGCTCTCTGACAGAGCCCTCCGAGAGCGGCGACGCGGGATCCGACGTTGATCGTAC 5880  
Db 6471 ACGGCTCTCTGACAGAGCCCTCCGAGAGCGGCGACGCGGGATCCGACGTTGATCGTAC 6530  
QY 5881 TCCTCATGCCCCCTTTCGAGGGGAGCGGGGATTCGCGATCTCAGCGAGCGGCTTTGG 5940  
Db 6531 TCCTCATGCCCCCTTTCGAGGGGAGCGGGGATTCGCGATCTCAGCGAGCGGCTTTGG 6590  
QY 5941 TCTAACGTAAGCGAGGAGCTAGTGTAGGAGCTGCTGCTGCTCGATGCTTCTACATG 6000  
Db 6591 TCTAACGTAAGCGAGGAGCTAGTGTAGGAGCTGCTGCTGCTCGATGCTTCTACATG 6050  
QY 6001 ACAGGCGCTGTATCAGCGATGCGTGTGGAGGAAACAGCTGCGCATCAATGACATG 6060  
Db 6651 ACAGGCGCTGTATCAGCGATGCGTGTGGAGGAAACAGCTGCGCATCAATGACATG 6120  
QY 6061 AGCAACTCTTGTCTCGTCCACCAACTTGTGTATGTACACATCTCGACGCGCAAGC 6120  
Db 6711 AGCAACTCTTGTCTCGTCCACCAACTTGTGTATGTGTACCAATCTCGACGCGCAAGC 6770  
QY 6121 CTGCGGACAGAAAGTCACTTTTGAAGTCTGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 6180  
Db 6771 CTGCGGACAGAAAGTCACTTTTGAAGTCTGAGTCTGAGGAGTCTGAGGAGTCTGAGGAG 6830  
QY 6181 GTGCTCAAGAGATGAAGCGGAGGCGTCAAGTGTAGGCTTAACTTCTATCCGTGGAG 6240  
Db 6831 GTGCTCAAGAGATGAAGCGGAGGCGTCAAGTGTAGGCTTAACTTCTATCCGTGGAG 6890  
QY 6241 GAAGCTGTAAAGTACGCGCCCACTTCGCGAGATCTAAATTTTGGCTATGGGCAAG 6300  
Db 6891 GAAGCTGTAAAGTACGCGCCCACTTCGCGAGATCTAAATTTTGGCTATGGGCAAG 6950  
QY 6301 GAGCTCGGAACTATTCAGCAAGGCGTTAAACCACTCCGCTCCGTGTGGAGGAGCTTG 6360  
Db 6951 GAGCTCGGAACTATTCAGCAAGGCGTTAAACCACTCCGCTCCGTGTGGAGGAGCTTG 7010  
QY 6361 CTGGAAGACACTGAGACACCAATTCACCACTATGAGGCTTAACTTCTATCCGTGGAG 6420  
Db 7011 CTGGAAGACACTGAGACACCAATTCACCACTATGAGGCTTAACTTCTATCCGTGGAG 7070  
QY 6421 GTCCACACAGAGAGGGGGCGGAGCGAGCTCGCTTATCGTATTTCCAGATTTGGG 6480  
Db 7071 GTCCACACAGAGAGGGGGCGGAGCGAGCTCGCTTATCGTATTTCCAGATTTGGG 7130  
QY 6481 GTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCTCCCTCAGGCGGTG 6540  
Db 7131 GTTCTGTGTGCGAGAAATGGCCCTTTACGATGTGGTCTCCACCTCTCCCTCAGGCGGTG 7190

QY	6541	ATGGGCTCTTTCATACGGATTCCAATACTCTCTGGACAGCGGTGCAAGTTCTCTGGTGAAT	6600
Db	7191	ATGGGCTCTTTCATACGGATTCCAATACTCTCTGGACAGCGGTGCAAGTTCTCTGGTGAAT	7250
QY	6601	GCCTGGAAAGGAAGAAATGCCCTATGGGCTTTCGCATATGACACACCGCTGTTTTGACTCA	6660
Db	7251	GCCTGGAAAGGAAGAAATGCCCTATGGGCTTTCGCATATGACACACCGCTGTTTTGACTCA	7310
QY	6661	ACGGTCACTGAGAAATGACATCCGTTGTGAGGAGTCAATCTACCAATGTTGTGACTTTGGCC	6720
Db	7311	ACGGTCACTGAGAAATGACATCCGTTGTGAGGAGTCAATCTACCAATGTTGTGACTTTGGCC	7370
QY	6721	CCGAAAGCCAGACAGGCGCATAAAGTTCGCTCACAGAGCGGCTTTACATCGGGGCCCCCTG	6780
Db	7371	CCGAAAGCCAGACAGGCGCATAAAGTTCGCTCACAGAGCGGCTTTACATCGGGGCCCCCTG	7430
QY	6781	ACTAAATCTTAAAGGCGAAGACTGCGGCTATCGCGGTGCGCGGAGCGGTGTACTGACG	6840
Db	7431	ACTAAATCTTAAAGGCGAAGACTGCGGCTATCGCGGTGCGCGGAGCGGTGTACTGACG	7490
QY	6841	ACCAAGCTGCGGTAAATACCCCTCACATGTTACTTGAAGGCCGCTCGCGGCTCTCGAGCTGCG	6900
Db	7491	ACCAAGCTGCGGTAAATACCCCTCACATGTTACTTGAAGGCCGCTCGCGGCTCTCGAGCTGCG	7550
QY	6901	AAGCTCCAGGACTGCACGATGCTCGTATGCGGAGACGACCTTGTCTTATCTGTGAAGC	6960
Db	7551	AAGCTCCAGGACTGCACGATGCTCGTATGCGGAGACGACCTTGTCTTATCTGTGAAGC	7610
QY	6961	GCGGGGACCCMAGAGGACGAGCGGAGCCTACGGGCTTTCACGAGGCTGTACTAGATAC	7020
Db	7611	GCGGGGACCCMAGAGGACGAGCGGAGCCTACGGGCTTTCACGAGGCTGTACTAGATAC	7670
QY	7021	TCTGCCCCCTCGGGGACCCGCCAAACACAGATAACGACTTGGAGTTGATAACATCATGC	7080
Db	7671	TCTGCCCCCTCGGGGACCCGCCAAACACAGATAACGACTTGGAGTTGATAACATCATGC	7730
QY	7081	TCCTCCAATGTGTCAAGTCGCGACAGTGCATCTGGCAAAAGGGTGTACTATCTCACCCGT	7140
Db	7731	TCCTCCAATGTGTCAAGTCGCGACAGTGCATCTGGCAAAAGGGTGTACTATCTCACCCGT	7790
QY	7141	GACCCGACACCCCTTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT	7200
Db	7791	GACCCGACACCCCTTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT	7850
QY	7201	TCCTGGCTTAGGCAACATCATCATGTATGCGCCACCTTGTGGCAAGGATGCTCTGATG	7260
Db	7851	TCCTGGCTTAGGCAACATCATCATGTATGCGCCACCTTGTGGCAAGGATGCTCTGATG	7910
QY	7261	ACTCATTTCTTCTCCATCTTCTAGCTCAGGAAACACTTGAAAAGCCCTAGATTGTCTAG	7320
Db	7911	ACTCATTTCTTCTCCATCTTCTAGCTCAGGAAACACTTGAAAAGCCCTAGATTGTCTAG	7970
QY	7321	ATCTACGGGGCTGTACTCATTTGAGGCACTTGACCTACCTCAGATCATTTCAACGACTC	7380
Db	7971	ATCTACGGGGCTGTACTCATTTGAGGCACTTGACCTACCTCAGATCATTTCAACGACTC	8030
QY	7381	CATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCT	7440
Db	8031	CATGGCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTGGCT	8090
QY	7441	TCATGCCCTCAGAAACTTGGGGTACCGCCCTTGCAGTCTGGAGACATCGGGGCCAAGT	7500
Db	8091	TCATGCCCTCAGAAACTTGGGGTACCGCCCTTGCAGTCTGGAGACATCGGGGCCAAGT	8150
QY	7501	GTCCGCGCTTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTTC	7560
Db	8151	GTCCGCGCTTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTTC	8210
QY	7561	AACCTGGGCAGTAAAGGACCAAGCTCAAATCACTCACTCCGCTCGCTCCAGTTGGAT	7620
Db	8211	AACCTGGGCAGTAAAGGACCAAGCTCAAATCACTCACTCCGCTCGCTCCAGTTGGAT	8270
QY	7621	TTATCCAGCTGGTTCGTTGCTGGTTTACAGCGGGGGAGACATATATACAGCTGTCTCGT	7680

Db	8271	TTATCCAGCTGGTTCGTTGCTGTGTTACAGCGGGGAGACATATATACAGCCCTGTCTCGT	8330
Qy	7681	GCCGAGCCCGCTGGTTCATGTGGTGCTACTCTACTTTCTCTAGGGGTAGGCATCTAT	7740
Db	8331	GCCGAGCCCGCTGGTTCATGTGGTGCTACTCTACTTTCTGTAGGGGTAGGCATCTAT	8390
Qy	7741	CTACTCCCCAACCGATGAACGGGAGCCTAAACACTCCAGGCCAATAGGCCATCTGTTTT	7800
Db	8391	CTACTCCCCAACCGATGAACGGGAGCCTAAACACTCCAGGCCAATAGGCCATCTGTTTT	8450
Qy	7801	TTTCCCTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCTTTT	7860
Db	8451	TTTCCCTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCTTTT	8510
Qy	7861	TTTTTTCCTCTTTTTTCTTTTTTCTTTTCTTTGTTGGTCTCCATCTTAGCCCTAGTCAAGGC	7920
Db	8511	TTTTTTCCTCTTTTTTCTTTTTTCTTTTCTTTGTTGGTCTCCATCTTAGCCCTAGTCAAGGC	8570
Qy	7921	TAGCTGTGAAGGTCCTGTAGCCGCTTGACTGCAGAGAGTGCTGATACTGGCCTCTCTGC	7980
Db	8571	TAGCTGTGAAGGTCCTGTAGCCGCTTGACTGCAGAGAGTGCTGATACTGGCCTCTCTGC	8630
Qy	7981	AGATCAAGT 7989	
Db	8631	AGATCAAGT 8639	
RESULT 9			
US-10-029-907-24			
; Sequence 24, Application US/10029907			
; Patent No. 6706874			
; GENERAL INFORMATION:			
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.			
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM			
; TITLE OF INVENTION: HEPATITIS C VIRUS			
; FILE REFERENCE: 13/063			
; CURRENT APPLICATION NUMBER: US/10/029,907			
; CURRENT FILING DATE: 2001-12-21			
; PRIOR APPLICATION NUMBER: 60/257,857			
; PRIOR FILING DATE: 2000-12-22			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 24			
; LENGTH: 8638			
; TYPE: DNA			
; ORGANISM: HCV			
; FEATURE:			
; NAME/KEY: CDS			
; LOCATION: (1802)...(8407)			
US-10-029-907-24			
Query Match 91.0%; Score 7271.2; DB 4; Length 8638;			
Best Local Similarity 92.2%; Pred. No. 0;			
Matches 7974; Conservative 0; Mismatches 3; Indels 671; Gaps 3;			
Qy	2	CCAGCCCCCGATTGGCGGCGACACTCCACCATAGATCACTCCCTCTGTGAGGAATACTCTGT	61
Db	2	CCAGCCCCCGATTGGCGGCGACACTCCACCATAGATCACTCCCTCTGTGAGGAATACTCTGT	61
Qy	62	CTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGTCGAGCGCTCCAGGACC	121
Db	62	CTTCACGAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGTGTCGAGCGCTCCAGGACC	121
Qy	122	CCCGCTCCGGGAGAGCATAGTGGTCTCGGAAACCGGTGAGTGATACCGGAATTGCCAGG	181
Db	122	CCCGCTCCGGGAGAGCATAGTGGTCTCGGAAACCGGTGAGTGATACCGGAATTGCCAGG	181
Qy	182	ACGACCGGTCCTTTCTTTGGATCAACCGGCTCAATGCTCGAGATTTGGGCGTGCCCGCG	241
Db	182	ACGACCGGTCCTTTCTTTGGATCAACCGGCTCAATGCTCGAGATTTGGGCGTGCCCGCG	241
Qy	242	CGAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAAGCGCTTGTGGTACTCTCCTCATAGGG	301



Db 242 CGAGACTGCTAGCGAGTAGTGTGGTTCGGAAGGCGCTTGTGTAATGCTGCTGATAGG 301  
QY 302 TGCTTGGAGTGGCCCGGAGGTCTCTAGACCGTGCACCATGAGCAGAACTCTAAACC 361  
Db 302 TGCTTGGAGTGGCCCGGAGGTCTCTAGACCGTGCACCATGAGCAGAACTCTAAACC 361  
QY 362 TCAAGAAACCAAGGCGCGCATGATTGAACAAGATGATGACCGAGTTCACCGAGTTCCTC 421  
Db 362 TCAAGAAACCAAGGCGCGCATGATTGAACAAGATGATGACCGAGTTCACCGAGTTCCTC 421  
QY 422 GGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAAATCGGCTGCTC 481  
Db 422 GGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAAGACAAATCGGCTGCTC 481  
QY 482 TGATGCCCGCGTGTTCGGCTGTACGGCAGGGGCGCCGGTCTTTTGTCAAGACCGA 541  
Db 482 TGATGCCCGCGTGTTCGGCTGTACGGCAGGGGCGCCGGTCTTTTGTCAAGACCGA 541  
QY 542 CTTGTCCGCTGCTGAAATGAATGCAAGTGCAGGACGAGCGCGGTATCGTGGCTGGCCAC 601  
Db 542 CTTGTCCGCTGCTGAAATGAATGCAAGTGCAGGACGAGCGCGGTATCGTGGCTGGCCAC 601  
QY 602 GACGGCGTTCCTTGGCAGCTGTGCTCGACGTGTCTCACTGAAGCGGGAAGGACTGGCT 661  
Db 602 GACGGCGTTCCTTGGCAGCTGTGCTCGACGTGTCTCACTGAAGCGGGAAGGACTGGCT 661  
QY 662 GCTATTGGGCGAAGTGGCCGGGAGGATCTCTGTCACTCACTGCTCTCTGCGGAGAA 721  
Db 662 GCTATTGGGCGAAGTGGCCGGGAGGATCTCTGTCACTCACTGCTCTCTGCGGAGAA 721  
QY 722 AGTATCCATCATGCTGATGCAATGCGCGCGGTGCAATACGCTTGATCGGCTACCTGCC 781  
Db 722 AGTATCCATCATGCTGATGCAATGCGCGCGGTGCAATACGCTTGATCGGCTACCTGCC 781  
QY 782 ATTTCGACCAACGAGAAACATCGCTCGAGCAGACGTAAGGATGGAAGCGGTCT 841  
Db 782 ATTTCGACCAACGAGAAACATCGCTCGAGCAGACGTAAGGATGGAAGCGGTCT 841  
QY 842 TGTGATCAGATGATCTGCAAGAGATCAGGGGCTCGCGCCAGCCGAACTGTTGCG 901  
Db 842 TGTGATCAGATGATCTGCAAGAGATCAGGGGCTCGCGCCAGCCGAACTGTTGCG 901  
QY 902 CAGGCTCAAGCGGCGATGCGCAGCGGAGGATCTGCTGACCATGGCGATGCTG 961  
Db 902 CAGGCTCAAGCGGCGATGCGCAGCGGAGGATCTGCTGACCATGGCGATGCTG 961  
QY 962 CTTGCCGAATATCATGTTGGAATATGCGCTTTTCTGGATTCATCGACTGTGGCGGCT 1021  
Db 962 CTTGCCGAATATCATGTTGGAATATGCGCTTTTCTGGATTCATCGACTGTGGCGGCT 1021  
QY 1022 GGTGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCGTGATATGCTGAAGACT 1081  
Db 1022 GGTGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCGTGATATGCTGAAGACT 1081  
QY 1082 TGGCGCGAATGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTCCCGATTCGCA 1141  
Db 1082 TGGCGCGAATGGCTGACCGCTTCTGCTGCTTTACGGTATCGCGCTCCCGATTCGCA 1141  
QY 1142 GCGCATCGCTTCTATCGCTTCTTGAAGAGTTCCTCTGAGTT-----TAAAC 1189  
Db 1142 GCGCATCGCTTCTATCGCTTCTTGAAGAGTTCCTCTGAGTTTCGCGCCAGATGTTAAC 1201  
QY 1190 AGACCAACCGTTTCCTTAGCGGATCAATTCGCGCCCTCTGCCCTCCCGCCCTTA 1249  
Db 1202 AGACCAACCGTTTCCTTAGCGGATCAATTCGCG-----CCCGCCCTTA 1250  
QY 1250 ACCTTACTGCGGACCGCTTGAATAAGCGCGGTGTGCTTGTCTATGTTATTTT 1309  
Db 1251 ACCTTACTGCGGACCGCTTGAATAAGCGCGGTGTGCTTGTCTATGTTATTTT 1310  
QY 1310 CCACCATATTGCGCTCTTTTGGCAATGTAGGCGCGGAACTGGCCCTGCTCTTGA 1369

Db 1311 CCACCATATTGCGCTCTTTTGGCAATGTAGGCGCGGAAACCTGGCCCTGCTCTTCTTGA 1370  
QY 1370 CGAGACTTCCTAGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTCG 1429  
Db 1371 CGAGACTTCCTAGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGTCG 1430  
QY 1430 TGAAGNAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTTAGCGACCTTTT 1489  
Db 1431 TGAAGNAGCAGTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTTAGCGACCTTTT 1490  
QY 1490 GCAGGACGCGAAACCCGCCACCTGGCGACAGTGCCTCTCGGCCAAAAGCACGTGTAT 1549  
Db 1491 GCAGGACGCGAAACCCGCCACCTGGCGACAGTGCCTCTCGGCCAAAAGCACGTGTAT 1550  
QY 1550 AAGATACCTCTCAAAAGGCGGCAACCCAGTGCACGTTGTGAGTTGATGTTGTTAGT 1609  
Db 1551 AAGATACCTCTCAAAAGGCGGCAACCCAGTGCACGTTGTGAGTTGATGTTGTTAGT 1610  
QY 1610 AAGAGTCAAAATGGCTCTCAAGCGTATTCACAAAGGGGCTGAAGATGCCAGAGG 1669  
Db 1611 AAGAGTCAAAATGGCTCTCAAGCGTATTCACAAAGGGGCTGAAGATGCCAGAGG 1670  
QY 1670 TACCCATTCTATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGTTTGTAGT 1729  
Db 1671 TACCCATTCTATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGTTTGTAGT 1730  
QY 1730 CGAGGTTAAAAACGCTTAGGCCCCCGAAACCAAGGGGACGTGTTTCTTTGAAAAAC 1789  
Db 1731 CGAGGTTAAAAACGCTTAGGCCCCCGAAACCAAGGGGACGTGTTTCTTTGAAAAAC 1790  
QY 1790 ACGATTAATCAATG----- 1803  
Db 1791 ACGATTAATCAATGACCGGAGATGCGACATCTGTCGGAGGCGGTTTTTCGTAGTCT 1850  
QY 1804 ----- 1803  
Db 1851 TGATACTCTTGACCTTGTACCCGACTATTAAGCTGTTCTCGTAGGCTCATATGTTGGT 1910  
QY 1804 ----- 1803  
Db 1911 TACAATATTTTATACACAGGCGGAGCACACTTGCAAGTGTGATCTCCCGCCCTCAACG 1970  
QY 1804 ----- 1803  
Db 1971 TTCCGGGGGCGCGATGCGTCACTCTCTCACGTGCGGATCCACCCAGAGCTAATCT 2030  
QY 1804 ----- 1803  
Db 2031 TTACCATCACCATAATCTTGTCTCGCCATCTCGGTCCACTCATGTTGCTCCAGGCTGGTA 2090  
QY 1804 ----- 1803  
Db 2091 TAACCAAGTGGCTGACTTCTGTCGGCGCACACGGGCTCATTCGTGCATGCTGTGTGC 2150  
QY 1804 ----- 1803  
Db 2151 GGAAGTGTGTCGGGCTCATTTATGTCCAAATGCTCTCATGAAGTTGGCGGCACTGACAG 2210  
QY 1804 ----- 1803  
Db 2211 GTACGTACGTTTATGACCATCTCACCCACTGGGACTGGGCCACGGCGGCTTAGGAG 2270  
QY 1804 ----- 1803  
Db 2271 ACCTTGGGTGGCAGTTAGCCCGTCTCTCTCTGATATGAGACCAAGTTATCACCT 2330  
QY 1804 ----- 1803  
Db 2331 GGGGGGAGACACCGCGGGTGTGGGACATCATCTTGGGCTGCGCGTCTCGCGCCGCA 2390  
QY 1804 ----- 1803  
Db 2391 GGGGGAGGAGATACATCTGGGACCGGACAGACGCTTGAAGGGGAGGGGTGGCGACTCC 2450

QY 1804 --GCGCCTTTACGGCCTACTCCCAACAGACGCGAGGCGCTACTTGGCTGCATCATCACTA 1861  
DB 2451 TCGCGCCTATTACGGCCTACTCCCAACAGACGCGAGGCGCTACTTGGCTGCATCATCACTA 2510  
QY 1862 GCCTCACAGGCGCGGACAGGAACCAAGTTCAGGGGAGGTCAAGTGGTCTCCACCGCAA 1921  
DB 2511 GCCTCACAGGCGCGGACAGGAACCAAGTTCAGGGGAGGTCAAGTGGTCTCCACCGCAA 2570  
QY 1922 CACAATCTTCTGTGGGACCTCGCTCAATGGCGTGTGTGGACTGTCTATCATGTGGCG 1981  
DB 2571 CACAATCTTCTGTGGGACCTCGCTCAATGGCGTGTGTGGACTGTCTATCATGTGGCG 2630  
QY 1982 GCTCAAGACCTTCCCGGCCCAAGGGCCCAATCACCAATGTACACCAATGTGGACC 2041  
DB 2631 GCTCAAGACCTTCCCGGCCCAAGGGCCCAATCACCAATGTACACCAATGTGGACC 2690  
QY 2042 AGGAACCTCGTCCGGCTGGCAAGCGCCCCCGGGCGCGTCTCTTGACACCATGCACTGG 2101  
DB 2691 AGGAACCTCGTCCGGCTGGCAAGCGCCCCCGGGCGCGTCTCTTGACACCATGCACTGG 2750  
QY 2102 GGAGCTCGGACCTTTACTTTGGTTCAGAGGATGCGGATGTCAATCCGGTGCSCCGGCGGG 2161  
DB 2751 GGAGCTCGGACCTTTACTTTGGTTCAGAGGATGCGGATGTCAATCCGGTGCSCCGGCGGG 2810  
QY 2162 GCGACGACGAGGGGAGGCTACTCTCCCGCAGGCGCGTCTCTACTTGAAGGGCTCTTCGG 2221  
DB 2811 GCGACGACGAGGGGAGGCTACTCTCCCGCAGGCGCGTCTCTACTTGAAGGGCTCTTCGG 2870  
QY 2222 GGGTCCACTGTCTGCCCCCTCGGGGACGCTGTGGGCATCTTTGGGCTGCCGTGTGCA 2281  
DB 2871 GGGTCCACTGTCTGCCCCCTCGGGGACGCTGTGGGCATCTTTGGGCTGCCGTGTGCA 2930  
QY 2282 CCGGAGGGGTTCGGAAGCGGTGGACTTTGTACCCGTGAGTCTATFGAAGCACTATGCG 2341  
DB 2931 CCGGAGGGGTTCGGAAGCGGTGGACTTTGTACCCGTGAGTCTATFGAAGCACTATGCG 2990  
QY 2342 GGTCCCGGTCTTACGGCAACTCGTCCCGCTCGGCGTACCGGAGCACTTCCAGGTGG 2401  
DB 2991 GGTCCCGGTCTTACGGCAACTCGTCCCGCTCGGCGTACCGGAGCACTTCCAGGTGG 3050  
QY 2402 CCATCTACACGCCCCCTACTGGTAGCGCAAGACTAAGGTGCGGGCTGGGTATGCA 2461  
DB 3051 CCATCTACACGCCCCCTACTGGTAGCGCAAGACTAAGGTGCGGGCTGGGTATGCA 3110  
QY 2462 CCAAGGTTAAAGGTGCTTGTCTGAAACCGGTTCGCGCCGCACTTCCAGGTGG 2521  
DB 3111 CCAAGGTTAAAGGTGCTTGTCTGAAACCGGTTCGCGCCGCACTTCCAGGTGG 3170  
QY 2522 CGTATATGTAAAGCACATGTATCGACCTAAACATCAGAACCGGGTAAGCACTCA 2581  
DB 3171 CGTATATGTAAAGCACATGTATCGACCTAAACATCAGAACCGGGTAAGCACTCA 3230  
QY 2582 CCACGGGTCCCCCATCACTCACTCCACCTATGGCAAGTTTCTTCCGCGCGGTGGTGTCT 2641  
DB 3231 CCACGGGTCCCCCATCACTCACTCCACCTATGGCAAGTTTCTTCCGCGCGGTGGTGTCT 3290  
QY 2642 CTGGGGGCGCCTATGACATCAATAATGTATGATGAGTGCCACTCAACTGACTGACCACTA 2701  
DB 3291 CTGGGGGCGCCTATGACATCAATAATGTATGATGAGTGCCACTCAACTGACTGACCACTA 3350  
QY 2702 TCCCTGGGATCGGCACAGTCTTGGACCAAGCGGAGAGCGCTGGAGCGGCACTCGTGTGC 2761  
DB 3351 TCCCTGGGATCGGCACAGTCTTGGACCAAGCGGAGAGCGCTGGAGCGGCACTCGTGTGC 3410  
QY 2762 TCGCCACCGCTACGCTCCGGGATCGGTTCACCGTGCCACATCCAAACATCGAGGAGGTGG 2821  
DB 3411 TCGCCACCGCTACGCTCCGGGATCGGTTCACCGTGCCACATCCAAACATCGAGGAGGTGG 3470  
QY 2822 CTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGGCATCCCGCATCGAGACCATCA 2881  
DB 3471 CTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGGCATCCCGCATCGAGACCATCA 3530

QY 2882 AGGGGGGAGGACCTCCTATTTCTGCCATTTCCAGAAGAAATGTGATGAGCTCGCGCGA 2941  
DB 3531 AGGGGGGAGGACCTCCTATTTCTGCCATTTCCAGAAGAAATGTGATGAGCTCGCGCGA 3590  
QY 2942 AGCTTCGGGCTCGGACTCAATGCTGTAGCATATTAACGGGCGCTTGAATGATCGTCA 3001  
DB 3591 AGCTTCGGGCTCGGACTCAATGCTGTAGCATATTAACGGGCGCTTGAATGATCGTCA 3650  
QY 3002 TACCAACTAGCGGAGACGTCAATGCTGTAGCAACCGAGCTCTAATAGCGGCTTTACCG 3061  
DB 3651 TACCAACTAGCGGAGACGTCAATGCTGTAGCAACCGAGCTCTAATAGCGGCTTTACCG 3710  
QY 3062 GGGATTTCCAGTCACTGATCGACTCAATGCTGTACCCAGACAGTCCGACTTCAGCC 3121  
DB 3711 GGGATTTCCAGTCACTGATCGACTCAATGCTGTACCCAGACAGTCCGACTTCAGCC 3770  
QY 3122 TGGACCCCACTTCACCAATTGAGACGACCGTGCACAAAGACGCGGTGTCACTGCTGC 3181  
DB 3771 TGGACCCCACTTCACCAATTGAGACGACCGTGCACAAAGACGCGGTGTCACTGCTGC 3830  
QY 3182 AGCGGGGAGGACGACTGGTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGGAG 3241  
DB 3831 AGCGGGGAGGACGACTGGTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGGAG 3890  
QY 3242 AACGGCCCTCGGCGCATGTTCCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGGTGTG 3301  
DB 3891 AACGGCCCTCGGCGCATGTTCCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGGTGTG 3950  
QY 3302 TTGTGTACGAGTCAACGCGCGGAGACCTCAAGTTAGTGTGCGGTGTACCTTAAACACAC 3361  
DB 3951 TTGTGTACGAGTCAACGCGCGGAGACCTCAAGTTAGTGTGCGGTGTACCTTAAACACAC 4010  
QY 3362 GAGGTTTCCCGTCTGCCAGGACCATCTGGAGTCTGGGAGAGCGCTTTTACAGGCTCA 3421  
DB 4011 GAGGTTTCCCGTCTGCCAGGACCATCTGGAGTCTGGGAGAGCGCTTTTACAGGCTCA 4070  
QY 3422 CCACATAGACGCCCATTTCTTGTCCCAGACTAAGCAGGACGAGACACTTCCCCTACC 3481  
DB 4071 CCACATAGACGCCCATTTCTTGTCCCAGACTAAGCAGGACGAGACACTTCCCCTACC 4130  
QY 3482 TGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGAGCC 3541  
DB 4131 TGGTAGCATACAGGCTACGGTGTGCGCAGGGCTCAGGCTCCACTCCATCGTGGAGCC 4190  
QY 3542 AAATGTGGAAGTGTCTCATACGCTAAAGCTTACGCTGACGCGGCCAACGCCCTGCTGT 3601  
DB 4191 AAATGTGGAAGTGTCTCATACGCTAAAGCTTACGCTGACGCGGCCAACGCCCTGCTGT 4250  
QY 3602 ATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACCAACACCCCATFAACCAATAATCA 3661  
DB 4251 ATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACCAACACCCCATFAACCAATAATCA 4310  
QY 3662 TGGCATGCAATGTCCGCTGACCTGAGGTGTGTCAAGGACCTTGGGTGTGTGAGCGGAG 3721  
DB 4311 TGGCATGCAATGTCCGCTGACCTGAGGTGTGTCAAGGACCTTGGGTGTGTGAGCGGAG 4370  
QY 3722 TCCTAGAGCTCTGCGCGGTATTCGTGACAAACAGGACGCGTGTGCTTGTGGGAGCA 3781  
DB 4371 TCCTAGAGCTCTGCGCGGTATTCGTGACAAACAGGACGCGTGTGCTTGTGGGAGCA 4430  
QY 3782 TCATCTTGTCCGAAAGCGCGGCATCATTTCCCGACAGGGAAGTCTCTTTACCGGGAGTTCG 3841  
DB 4431 TCATCTTGTCCGAAAGCGCGGCATCATTTCCCGACAGGGAAGTCTCTTTACCGGGAGTTCG 4490  
QY 3842 ATGAGATGGAAGAGTGTGCGCTCACCTCCCTTATCATGAAACAGGGAATGCAAGTTCGCCG 3901  
DB 4491 ATGAGATGGAAGAGTGTGCGCTCACCTCCCTTATCATGAAACAGGGAATGCAAGTTCGCCG 4550  
QY 3902 AACAAATCAAAACAGAGGCAATCGGTTCTGCAAAACGCCACCAAGCAAGCGAGGCTG 3961  
DB 4551 AACAAATCAAAACAGAGGCAATCGGTTCTGCAAAACGCCACCAAGCAAGCGAGGCTG 4610  
QY 3962 CTGCTCCGCTGTGGAATCCAAAGTGGCGGACCCCTCGAAGCCCTTCTGGGGAAGCATATGT 4021



QY 6182 TGCTCAAGGAGATGAAGCGGAGCGGTCCACAGTTAAAGCTAAACTTCTATCCGTGGAGG 6241  
Db 6831 TGCTCAAGGAGATGAAGCGGAGCGGTCCACAGTTAAAGCTAAACTTCTATCCGTGGAGG 6890  
QY 6242 AAGCCTGTAAAGTGAAGCGGCGCCACACATTCGGCCAGATCTAAATTTTGGCTATGGGCGAAAGG 6301  
Db 6891 AAGCCTGTAAAGTGAAGCGGCGCCACACATTCGGCCAGATCTAAATTTTGGCTATGGGCGAAAGG 6950  
QY 6302 AGTCCGGAACTATTCAGAGAGCGCGTTAAACCAATCCGCTCCGTGTGGAAAGCACTTGC 6361  
Db 6951 AGTCCGGAACTATTCAGAGAGCGCGTTAAACCAATCCGCTCCGTGTGGAAAGCACTTGC 7010  
QY 6362 TGGAAAGACACTGAGACACCAATTGACACACCATCATGCAAAATAAGGTTTCTCGG 6421  
Db 7011 TGGAAAGACACTGAGACACCAATTGACACACCATCATGCAAAATAAGGTTTCTCGG 7070  
QY 6422 TCCAAACAGAGAGAGGCGCGCAAGCCAGCTCGCCTTATCGTATTCACAGATTTGGGG 6481  
Db 7071 TCCAAACAGAGAGAGGCGCGCAAGCCAGCTCGCCTTATCGTATTCACAGATTTGGGG 7130  
QY 6482 TTCGTGTGCGAGAAATGGCCCTTTAGGATGCTCTCACCTCCCTCAGGCCGTGA 6541  
Db 7131 TTCGTGTGCGAGAAATGGCCCTTTAGGATGCTCTCACCTCCCTCAGGCCGTGA 7190  
QY 6542 TGGGCTCTTCATACGGATTCCAAATCTCTCTGGACAGCGGTCCAGTTCCTGGTGAATG 6601  
Db 7191 TGGGCTCTTCATACGGATTCCAAATCTCTCTGGACAGCGGTCCAGTTCCTGGTGAATG 7250  
QY 6602 COTGAAAGCGAAGAAATGCCCTATGGCTTCGCATATGACACCGCGCTGTTTGACTCAA 6661  
Db 7251 COTGAAAGCGAAGAAATGCCCTATGGCTTCGCATATGACACCGCGCTGTTTGACTCAA 7310  
QY 6662 CGGTCACTGAGAAATGACATCGTGTGAGAGTCAATCTACCAATGTTGACTTGGGCC 6721  
Db 7311 CGGTCACTGAGAAATGACATCGTGTGAGAGTCAATCTACCAATGTTGACTTGGGCC 7370  
QY 6722 CCGAAGCGAGACAGGCCATAAGTTCGCTCACAGAGCGGCTTACATCGGGGCGCCCTGA 6781  
Db 7371 CCGAAGCGAGACAGGCCATAAGTTCGCTCACAGAGCGGCTTACATCGGGGCGCCCTGA 7430  
QY 6782 CTAATTTCAAAGGCGAGAACTCGGCTATCGCGGTGCGCGAGCGGTGTACTGACGA 6841  
Db 7431 CTAATTTCAAAGGCGAGAACTCGGCTATCGCGGTGCGCGAGCGGTGTACTGACGA 7490  
QY 6842 CCAGTGGGCTAAATCCCTCACAATGTTACTTGAAGCGCGCTGCGGCCCTGTGAGCTGCGA 6901  
Db 7491 CCAGTGGGCTAAATCCCTCACAATGTTACTTGAAGCGCGCTGCGGCCCTGTGAGCTGCGA 7550  
QY 6902 AGCTCCAGGACTGCACGATGCTCGTATCGGAGAGCGACCTTGTCTGTTATCTGTGAAGCG 6961  
Db 7551 AGCTCCAGGACTGCACGATGCTCGTATCGGAGAGCGACCTTGTCTGTTATCTGTGAAGCG 7610  
QY 6962 CGGGAGCCCAAGAGAGAGCGAGCGCTTACCGGCGCTTCACGAGGCTATGACTAGATACT 7021  
Db 7611 CGGGAGCCCAAGAGAGAGCGAGCGCTTACCGGCGCTTCACGAGGCTATGACTAGATACT 7670  
QY 7022 CTGCCCCCTTGGGACCGCGCCCAACACAGATACGACTTGGAGTTGATAATCATGCT 7081  
Db 7671 CTGCCCCCTTGGGACCGCGCCCAACACAGATACGACTTGGAGTTGATAATCATGCT 7730  
QY 7082 CCTCAATGTGTCACTCGCGCACGATGCTCTGGCAAAAGGCTGTACTATCTCACCCGTG 7141  
Db 7731 CCTCAATGTGTCACTCGCGCACGATGCTCTGGCAAAAGGCTGTACTATCTCACCCGTG 7790  
QY 7142 ACCCCACACCCCTTGGCGGGGTGCGTGGGAGACAGCTAGACACACTCCAGTCAATT 7201  
Db 7791 ACCCCACACCCCTTGGCGGGGTGCGTGGGAGACAGCTAGACACACTCCAGTCAATT 7850  
QY 7202 CCTGGTAGGCAACATCATGATGATGCGCCACCTTGTGGCAAGGATGATCCTGATGA 7261  
Db 7851 CCTGGTAGGCAACATCATGATGATGCGCCACCTTGTGGCAAGGATGATCCTGATGA 7910

QY 7262 CTCATTTCTTCTCCATCTCTTAGCTCAGGAACAACCTTGAAAAAGCCCTAGATTGTGACA 7321  
Db 7911 CTCATTTCTTCTCCATCTCTTAGCTCAGGAACAACCTTGAAAAAGCCCTAGATTGTGACA 7970  
QY 7322 TCTACGGGCGCTGTTTACTCCATTGAGCCACTTGACTACCTCAGATCAATTCAACGACTCC 7381  
Db 7971 TCTACGGGCGCTGTTTACTCCATTGAGCCACTTGACTACCTCAGATCAATTCAACGACTCC 8030  
QY 7382 ATGGCCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCTT 7441  
Db 8031 ACGGCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGGTGGCTT 8090  
QY 7442 CATGCTCAGAAACTTCGGGTACCGCTTCGAGTCTGGAGACATCGGGCCAGAAGTG 7501  
Db 8091 CATGCTCAGAAACTTCGGGTACCGCTTCGAGTCTGGAGACATCGGGCCAGAAGTG 8150  
QY 7502 TCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTCA 7561  
Db 8151 TCCGCGCTAGGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTCA 8210  
QY 7562 ACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCCAATCCGGCTCGGTCCCAGTTGGATT 7621  
Db 8211 ACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCCAATCCGGCTCGGTCCCAGTTGGATT 8270  
QY 7622 TATCCAGCTGTGTTGCTGCTTACAGCGGGGAGACATATATCACAGCCTGTCTCGTG 7681  
Db 8271 TATCCAGCTGTGTTGCTGCTTACAGCGGGGAGACATATATCACAGCCTGTCTCGTG 8330  
QY 7682 CCGACCCCGCTGTTCAATGTGGTGCCTACTCTCTACTTTCTGTAGGGGTAGGCATCTATC 7741  
Db 8331 CCGACCCCGCTGTTCAATGTGGTGCCTACTCTCTACTTTCTGTAGGGGTAGGCATCTATC 8390  
QY 7742 TACTCCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAAATAGGCCATCTCTGTTTT 7801  
Db 8391 TACTCCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAAATAGGCCATCTCTGTTTT 8450  
QY 7802 TTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7861  
Db 8451 TTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 8510  
QY 7862 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7921  
Db 8511 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 8570  
QY 7922 AGCTGTGAAGCTCGGTGAGCGCTTGACTGAGAGAGTGTGATGACTGAGCTCTGCA 7981  
Db 8571 AGCTGTGAAGCTCGGTGAGCGCTTGACTGAGAGAGTGTGATGACTGAGCTCTGCA 8630  
QY 7982 GATCAAGT 7989  
Db 8631 GATCAAGT 8638

## RESULT 10

US-10-029-907-7  
; Sequence 7, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; CURRENT FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 7  
; LENGTH: 8638  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:

NAME/KEY: CDS  
LOCATION: (1802)...(8407)  
US-10-029-907-7

Query Match 90.7%; Score 7248.2; DB 4; Length 8638;  
Best Local Similarity 92.0%; Pred. No. 0;  
Matches 7960; Conservative 0; Mismatches 18; Indels 671; Gaps 3;

QY 1 GCCAGCCCCGATGGGGGGGACACTCCACATAGATCACTCCCTCTGTGAGAACTACTG 60  
Db 1 GCCAGCCCCGATGGGGGGGACACTCCACATAGATCACTCCCTCTGTGAGAACTACTG 60

QY 61 TCTTACGACAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGCNGCAGCCTCCAGGAC 120  
Db 61 TCTTACGACAGAAAGCGTCTAGCCATGGCGTTAGTATGAGTGCNGCAGCCTCCAGGAC 120

QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180  
Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180

QY 181 GAGACCGGGTCTTTCTTGGATCAACCGCTCAATGCTGAGATTTGGGGGTGCCCCC 240  
Db 181 GAGACCGGGTCTTTCTTGGATCAACCGCTCAATGCTGAGATTTGGGGGTGCCCCC 240

QY 241 GCGAGACTGCTAGCGAGTAGTGTGGGTGCGGAAAGGCTTGTGTAATGCTGATAGG 300  
Db 241 GCGAGACTGCTAGCGAGTAGTGTGGGTGCGGAAAGGCTTGTGTAATGCTGATAGG 300

QY 301 GTGCTTGCAGTGCCTCCCGGAGGCTCTGTAGACCGTGCACCATGAGCAGCAATCTTAAAC 360  
Db 301 GTGCTTGCAGTGCCTCCCGGAGGCTCTGTAGACCGTGCACCATGAGCAGCAATCTTAAAC 360

QY 361 CTCAGAGAAACCAAGAGCGGCGCATGATGAAACAGATGATGACAGCGAGTTCTC 420  
Db 361 CTCAGAGAAACCAAGAGCGGCGCATGATGAAACAGATGATGACAGCGAGTTCTC 420

QY 421 CGGCGCTTGGGTGGAGAGGCTATTGGGTATGACTGGGCAACAGCAATCGGCTGT 480  
Db 421 CGGCGCTTGGGTGGAGAGGCTATTGGGTATGACTGGGCAACAGCAATCGGCTGT 480

QY 481 CTGATGCCCGGTGTTCCGGCTGTACAGCGAGGGGCGCCGGTCTTTTGTCAAGACCG 540  
Db 481 CTGATGCCCGGTGTTCCGGCTGTACAGCGAGGGGCGCCGGTCTTTTGTCAAGACCG 540

QY 541 ACTGTCCGGTCCCTGATGAACTGACAGAGAGCGGCGGTATCGTGGCTGGCCA 600  
Db 541 ACTGTCCGGTCCCTGATGAACTGACAGAGAGCGGCGGTATCGTGGCTGGCCA 600

QY 601 CGACGGCGCTTCTTCGCGAGCTGTCTGACGTTGTCACTGAAGCGGGAAGGACTGGC 660  
Db 601 CGACGGCGCTTCTTCGCGAGCTGTCTGACGTTGTCACTGAAGCGGGAAGGACTGGC 660

QY 661 TGTATTGGGCGAAGTCCCGGGGAGGATCTCTGTCACTCACTTGTCTTCCGCGAGA 720  
Db 661 TGTATTGGGCGAAGTCCCGGGGAGGATCTCTGTCACTCACTTGTCTTCCGCGAGA 720

QY 721 AGTATCCATATGGCTGATGCAATGGCGGCTGTGATACGCTGTGATCGGCTGCTGCC 780  
Db 721 AGTATCCATATGGCTGATGCAATGGCGGCTGTGATACGCTGTGATCGGCTGCTGCC 780

QY 781 CATTCGACCAACGAGCAACATCGCATCGAGGAGCAGTACTTCGATGGAGCGGTC 840  
Db 781 CATTCGACCAACGAGCAACATCGCATCGAGGAGCAGTACTTCGATGGAGCGGTC 840

QY 841 TTGTCGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGGCGCAGCGAACTGTTCG 900  
Db 841 TTGTCGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGGCGCAGCGAACTGTTCG 900

QY 901 CCAGGCTCAAGCGCGCATGCCGAGCGGAGGATCTCGTGTGACCCATGGCGATGCCT 960  
Db 901 CCAGGCTCAAGCGCGCATGCCGAGCGGAGGATCTCGTGTGACCCATGGCGATGCCT 960

QY 961 GCTTGGCGAATATCATGTTGGAAATGGCGCTTTTCTGGAATTCATCGACTGTGGCGGC 1020

Db 961 GCTTGGCGAATATCATGTTGGAAATGGCGCTTTTCTGGAATTCATCGACTGTGGCGGC 1020

QY 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATGCTGAAGAGC 1080  
Db 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGCGTTGGCTACCGGTGATATGCTGAAGAGC 1080

QY 1081 TTGGCGGCGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCGCTCCCGATTCCG 1140  
Db 1081 TTGGCGGCGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCGCTCCCGATTCCG 1140

QY 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTCGAGTT-----TAA 1188  
Db 1141 AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTCGAGTTCGCGGCCAGATGTAA 1200

QY 1189 CAGACCAACCGTTCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCCCT 1248  
Db 1201 CAGACCAACCGTTCCTCTAGCGGGATCAATTCG-----CCCCCCCCCT 1249

QY 1249 AAGCTTACTGGCGGAGCGCTTGGATTAAGCGCGGTGCGGTGCTATATGTTATTT 1308  
Db 1250 AAGCTTACTGGCGGAGCGCTTGGATTAAGCGCGGTGCGGTGCTATATGTTATTT 1309

QY 1309 TCCACCATATTCGCGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGTCTCTTG 1368  
Db 1310 TCCACCATATTCGCGTCTTTTGGCAATGTAGGGCCCGGAAACCTGGCCCTGTCTCTTG 1369

QY 1369 ACAGCATTCCTAGGGGTCTTCCCTCTCGCAAGAGAAATGCAAGGTCTGTGAAATGTC 1428  
Db 1370 ACAGCATTCCTAGGGGTCTTCCCTCTCGCAAGAGAAATGCAAGGTCTGTGAAATGTC 1429

QY 1429 GTGAGGAGCAGTTCCTCTGGAAGTCTTGAAGCAACACACCTCTGTAGCGACCTT 1488  
Db 1430 GTGAGGAGCAGTTCCTCTGGAAGTCTTGAAGCAACACACCTCTGTAGCGACCTT 1489

QY 1489 TGCAGGACGCGAAACCCCGACCTGCGACAGGTGCTCTGCGGCCAAAAGCACCTGTA 1548  
Db 1490 TGCAGGACGCGAAACCCCGACCTGCGACAGGTGCTCTGCGGCCAAAAGCACCTGTA 1549

QY 1549 TAAGATACCTGTCAAGAGCGGCAACCCCGACCTGCGACAGGTGCTCTGCGGCCAAAAGCACCTGTA 1608  
Db 1550 TAAGATACCTGTCAAGAGCGGCAACCCCGACCTGCGACAGGTGCTCTGCGGCCAAAAGCACCTGTA 1609

QY 1609 GAAAGTCAATGCTCTCTCAAGCTATTCAACAAAGGGCTGAAGATGCCAGAAG 1668  
Db 1610 GAAAGTCAATGCTCTCTCAAGCTATTCAACAAAGGGCTGAAGATGCCAGAAG 1669

QY 1669 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTAG 1728  
Db 1670 GTACCCCATTTGATGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTAG 1729

QY 1729 TCAGGTTAAAAACGTCTAGGCCCCCGAACAACCGGAGCGTGTGTTTCTTTGAAAAA 1788  
Db 1730 TCAGGTTAAAAACGTCTAGGCCCCCGAACAACCGGAGCGTGTGTTTCTTTGAAAAA 1789

QY 1789 CAGGATAATACATG----- 1803  
Db 1790 CAGGATAATACATG----- 1803

QY 1804 ----- 1803  
Db 1850 CTGATCTCTGACCTTGTACCGCACTATAAGCTGTCTCGTAGGCTCATATGTTGG 1909

QY 1804 ----- 1803  
Db 1910 TTACATAATTTTATCACCAGGGCCGAGGACACACTTGAAGTGTGGATCCCCCCCCCTCAAC 1969

QY 1804 ----- 1803  
Db 1970 GTTCGGGGGGCGCGATGCCGTCTCTCTCAGTGGCGGATCCACCCAGAGCTAATC 2029

QY 1804 ----- 1803

Db 2030 TTTACCATCAGCAAAATCTTGCTGCCATCTCGGTCCACTCATGGTGTCCAGGCTGGT 2089  
Qy 1804 ----- 1803  
Db 2090 ATAACCAAGTGCCGTACTTGCTGCGCGACACGGGCTCATTCGTGCATGATGTGTG 2149  
Qy 1804 ----- 1803  
Db 2150 CGGAAGTTGCTGGGGTCAATATGTCCAAATGGCTCTCATGAAGTTGGCCGCACTGACA 2209  
Qy 1804 ----- 1803  
Db 2210 GGTACGTACTGTTATGAOCATCTCACCCCACTGCGGGAGCTGGGCCACCGCGGGCTACGA 2269  
Qy 1804 ----- 1803  
Db 2270 GACCTTGGGTGGCAGTTGAGCCCGTGTCTTCTGTGATATGGAGACCAAGTTATCACC 2329  
Qy 1804 ----- 1803  
Db 2330 TGGGGGCGAGACACCGGCGGTGTGGGACATCATCTTGGGCTGCCCGTCTCCGCCGCG 2389  
Qy 1804 ----- 1803  
Db 2390 AGGGGAGGGAGATACATCTGGGACCGGCGAGACGCTTGAAGGCGAGGGGTGGCGACTC 2449  
Qy 1804 ---GCGCCTATTACGGCTACTCCCAACAGAGCGAGCGCTACTTTGGCTGCATCATCT 1860  
Db 2450 CTCGGCCCTATTACGGCTACTCCCAACAGAGCGAGCGCTACTTTGGCTGCATCATCTACC 2509  
Qy 1861 AGCCTCACAGGCGGGGACAGGAACAGAGTTCAGGGGGAGGTCCTCAAGTGTCTTCCACCGCA 1920  
Db 2510 AGCCTCACAGGCGGGGACAGGAACAGAGTTCAGGGGGAGGTCCTCAAGTGTCTTCCACCGCA 2569  
Qy 1921 ACACAACTTTCCTGGGACCTGCTGCTAAATGGCGTGTGGAGCTGTCTATCATGGTGGC 1980  
Db 2570 ACACAACTTTCCTGGGACCTGCTGCTAAATGGCGTGTGGAGCTGTCTATCATGGTGGC 2629  
Qy 1981 GGCTCAAGAGCCCTTGGCGGCCCAAGAGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
Db 2630 GGCTCAAGAGCCCTTGGCGGCCCAAGAGGCCCAATCACCCAAATGTACCAATGTGGAC 2689  
Qy 2041 CAGGACCTCGTGGGTGCAAGCGCCCGCGGGCGCTTCTTGACACCACTGACCTGCG 2100  
Db 2690 CAGGACCTCGTGGGTGCAAGCGCCCGCGGGCGCTTCTTGACACCACTGACCTGCG 2749  
Qy 2101 GCGAGCTCGGACCTTTACTTGGTCAAGAGGATGCCGATGTCAATCCGGTGGCGCGCG 2160  
Db 2750 GCGAGCTCGGACCTTTACTTGGTCAAGAGGATGCCGATGTCAATCCGGTGGCGCGCG 2809  
Qy 2161 GCGGACAGCGGGGAGCTACTCTCCCGCGAGCCCGCTCTCTACTTGAAGGCTCTTCG 2220  
Db 2810 GCGGACAGCGGGGAGCTACTCTCCCGCGAGCCCGCTCTCTACTTGAAGGCTCTTCG 2869  
Qy 2221 GCGGCTCOACTGCTCTGCGCCCTCGGGGCACTGTGGGCACTTTTCGGGCTGCCGTGTGC 2280  
Db 2870 GCGGCTCOACTGCTCTGCGCCCTCGGGGCACTGTGGGCACTTTTCGGGCTGCCGTGTGC 2929  
Qy 2281 ACCGAGGGGTGGAGGGGTGACATTTGTATCCGCTCGAGTCTATGGAAACCACTATG 2340  
Db 2930 ACCGAGGGGTGGAGGGGTGACATTTGTATCCGCTCGAGTCTATGGAAACCACTATG 2989  
Qy 2341 CGGTCCCGGCTTTACAGGACAACTCGTCCCTCCGGCGGTACCGGACATTCAGGTG 2400  
Db 2990 CGGTCCCGGCTTTACAGGACAACTCGTCCCTCCGGCGGTACCGGACATTCAGGTG 3049  
Qy 2401 GCCCATCTACAGCCCTTACTGGTAGCGGCAAGAGCACTAAGGTGCCGGCTGCGTATGCA 2460  
Db 3050 GCCCATCTACAGCCCTTACTGGTAGCGGCAAGAGCACTAAGGTGCCGGCTGCGTATGCA 3109  
Qy 2461 GCCGAGGTATAGGTGCTTGTCTGAAACCGGCTCGCGGCAACCTTAGGTTCGGG 2520  
Db 3110 GCCGAGGTATAGGTGCTTGTCTGAAACCGGCTCGCGGCAACCTTAGGTTCGGG 3169

Qy 2521 GCGTATATGTCTAAGGCACATGTTATCGACCTTAACATCAGAAACCGGGTAAGGACCATC 2580  
Db 3170 GCGTATATGTCTAAGGCACATGTTATCGACCTTAACATCAGAAACCGGGTAAGGACCATC 3229  
Qy 2581 ACCAGGGTGGCCCGCATCACGTACTTCCACCTATGGCAAGTTTCTTTCGCGAGCGTGTGTC 2640  
Db 3230 ACCAGGGTGGCCCGCATCACGTACTTCCACCTATGGCAAGTTTCTTTCGCGAGCGTGTGTC 3289  
Qy 2641 TCTGGGGCGGCTATGACATCATATATGTGATGATGCCACTAATCTGACTCGACCACT 2700  
Db 3290 TCTGGGGCGGCTATGACATCATATATGTGATGATGCCACTAATCTGACTCGACCACT 3349  
Qy 2701 ATCTTGGGCATCGGCACAGTCTCTGACCAAGCGAGCGGTGGAGCGGCTCTCTGTCG 2760  
Db 3350 ATCTTGGGCATCGGCACAGTCTCTGACCAAGCGAGCGGTGGAGCGGCTCTCTGTCG 3409  
Qy 2761 CTCGCCACCGTACGCTTCCGGGATCGGTCCCGTCCACATCCAAACATCGAGGAGGTG 2820  
Db 3410 CTCGCCACCGTACGCTTCCGGGATCGGTCCCGTCCACATCCAAACATCGAGGAGGTG 3469  
Qy 2821 GCTCTGTCAGGACTGGAGAAATCCCTTTTATGGCAAGGCAATCCCATCGAGACCATC 2880  
Db 3470 GCTCTGTCAGGACTGGAGAAATCCCTTTTATGGCAAGGCAATCCCATCGAGACCATC 3529  
Qy 2881 AAGGGGGGAGGCACCTCATTTCTGCCATTCRAAGAGAAATGTGATGAGTCCCGCG 2940  
Db 3530 AAGGGGGGAGGCACCTCATTTCTGCCATTCRAAGAGAAATGTGATGAGTCCCGCG 3589  
Qy 2941 AAGTGTTCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATGATCGTGC 3000  
Db 3590 AAGTGTTCGGCTCGGACTCAATGCTGTAGCATATTAACGGGGCTTGTATGATCGTGC 3649  
Qy 3001 ATACCACTAGCGGAGAGCTCATTTGCTGAGCAAGGAGCGCTCTAATCAAGCGGCTTACC 3060  
Db 3650 ATACCACTAGCGGAGAGCTCATTTGCTGAGCAAGGAGCGCTCTAATCAAGCGGCTTACC 3709  
Qy 3061 GCGGATTCGACTCAGTGATGACTGCAATACATGTGTCAACCAGACAGTGCATTTAGG 3120  
Db 3710 GCGGATTCGACTCAGTGATGACTGCAATACATGTGTCAACCAGACAGTGCATTTAGG 3769  
Qy 3121 CTGACCCGACCTTTACCAATGAGACGACGCGTCCACAGAGCGGCTGTCAAGCTCG 3180  
Db 3770 CTGACCCGACCTTTACCAATGAGACGACGCGTCCACAGAGCGGCTGTCAAGCTCG 3829  
Qy 3181 CAGCGGCGAGCGAGGACTGGTAGGGGAGGATGGCATTTACAGTTTGTGACTCCAGG 3240  
Db 3830 CAGCGGCGAGCGAGGACTGGTAGGGGAGGATGGCATTTACAGTTTGTGACTCCAGG 3889  
Qy 3241 GAACGGCCCTCGGCACTGTTCTGCTCGGTTCTGTCGAGTGTCTATGACGCGGCTGT 3300  
Db 3890 GAACGGCCCTCGGCACTGTTCTGCTCGGTTCTGTCGAGTGTCTATGACGCGGCTGT 3949  
Qy 3301 GCTTGTACGAGCTCAGCCCGCGGAGACCTCAGTTAGTGTTCGGGCTTACTTAACACA 3360  
Db 3950 GCTTGTACGAGCTCAGCCCGCGGAGACCTCAGTTAGTGTTCGGGCTTACTTAACACA 4009  
Qy 3361 CAGGGTTGCCGCTTCCAGGACCACTGAGTTCTGGGAGAGGCTTTTACAGGCTC 3420  
Db 4010 CAGGGTTGCCGCTTCCAGGACCACTGAGTTCTGGGAGGCGCTTTTACAGGCTC 4069  
Qy 3421 ACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGCGAGAGCAACTTTCCTTAC 3480  
Db 4070 ACCCACATAGACGCCCATTTCTTGTCCAGACTAAGCAGGCGAGAGCAACTTTCCTTAC 4129  
Qy 3481 CTGTAGCATACAGGCTACGGTGTGCGCAGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
Db 4130 CTGTAGCATACAGGCTACGGTGTGCGCAGGCTCAGGCTCCACCTCCATCGTGGGAC 4189  
Qy 3541 CAATGTGGAGTGTCTCATACGCTAAGCCTACGCTGACGGGCGGCGGCGGCGGCGGCTG 3600  
Db 4190 CAATGTGGAGTGTCTCATACGCTAAGCCTACGCTGACGGGCGGCGGCGGCGGCGGCTG 4249



QY 3601 TATAGCTGGAGCGGTTCAAAACAGAGTTACTACACACACCCCATACCAATAATACATC 3660  
Db |||||  
QY 4250 TATAGCTGGAGCGGTTCAAAACAGAGTTACTACACACACCCCATACCAATAATACATC 4309  
Db |||||  
QY 3661 ATGGATGATCATCGGCTGACCTGGAGTGTGTCACGAGCACCTGGTGTGTTAGGCGGA 3720  
Db |||||  
QY 4310 ATGGATGATCATCGGCTGACCTGGAGTGTGTCACGAGCACCTGGTGTGTTAGGCGGA 4369  
Db |||||  
QY 3721 GTCTAGAGCTCTGGCCGCGTATTGCTGTGAACACAGGAGCGGTGTGTATTTGGGCGAG 3780  
Db |||||  
QY 4370 GTCCCTAGCAGCTCTGGCTCGTATTGCTGTGACAAACAGGAGCGGTGTGTATTTGGGCGAG 4429  
QY 3781 ATCATCTCTCGGAAAGCCGCGCATATTCCCGACAGAGGAGTCTTTACCGGAGTTTC 3840  
Db |||||  
QY 4430 ATCACTCTGTCGGAAGCGCGCATATTCCCGACAGGAAAGTCTTTACCGGAGTTTC 4489  
Db |||||  
QY 3841 GATGATGAGGAGTGTGCTTACACCTCCCTTATCATCGAAACAGGGAATGACGCTCGCC 3900  
Db |||||  
QY 4490 GATGATGAGGAGTGTGCTTACACCTCCCTTATCATCGAAACAGGGAATGACGCTCGCC 4549  
QY 3901 GAACAAATTCAAACAGAGGCAATCGGCTGTGCAAAACAGGCAACAGGAGCGGT 3960  
Db |||||  
QY 4550 GAACAAATTCAAACAGAGGCAATCGGCTGTGCAAAACAGGCAACAGGAGCGGT 4609  
QY 3961 GCTGCTCCGCTGTGAATCCAGTGGCGGACCTCGAGGCTTCTGGGCGAGCATATG 4020  
Db |||||  
QY 4610 GCTGCTCCGCTGTGAATCCAGTGGCGGACCTCGAGGCTTCTGGGCGAGCATATG 4669  
QY 4021 TGGAAATTCATCAGCGGGATACAATATTTAGCAGGCTTGTCCACTCTGCTGTGCAACCC 4080  
Db |||||  
QY 4670 TGGAAATTCATCAGCGGGATACAATATTTAGCAGGCTTGTCCACTCTGCTGTGCAACCC 4729  
QY 4081 GCGATAGATCACTGATGAGATTCACAGCTCTATACACAGGCGGCTACACACCAACAT 4140  
Db |||||  
QY 4730 GCGATAGATCACTGATGAGATTCACAGCTCTATACACAGGCGGCTACACACCAACAT 4789  
QY 4141 ACCCTCTCTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTCTCTCCAGCGCT 4200  
Db |||||  
QY 4790 ACCCTCTCTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTCTCTCCAGCGCT 4849  
QY 4201 GCTTCTGCTTTTCTAGCGCGGCATCGCTGGAGCGGCTGTTGGACGATAGGCTTTGG 4260  
Db |||||  
QY 4850 GCTTCTGCTTTTCTAGCGCGGCATCGCTGGAGCGGCTGTTGGACGATAGGCTTTGG 4309  
QY 4261 AAGGTCTTGTGATATTTTGGCAGTTATGAGCAGGGGTGGCAGCGGCTCGTGGCC 4320  
Db |||||  
QY 4910 AAGGTCTTGTGATATTTTGGCAGTTATGAGCAGGGGTGGCAGCGGCTCGTGGCC 4369  
QY 4321 TTTAAGGTATGAGCGGCGAGATGCCCTCCACCGAGACCTGGTTAACTTACTTCCCTGCT 4380  
Db |||||  
QY 4970 TTTAAGGTATGAGCGGCGAGATGCCCTCCACCGAGACCTGGTTAACTTACTTCCCTGCT 5029  
QY 4381 ATCCCTCTCCCTTGGCGCCCTAGTCTGGGGTGTGTCGCGAGCATATGCGTCTGGGCAC 4440  
Db |||||  
QY 5030 ATCCCTCTCCCTTGGCGCCCTAGTCTGGGGTGTGTCGCGAGCATATGCGTCTGGGCAC 5089  
QY 4441 GTGGGCGGAGGGGCTGTGAGTGTGATGAACCGGCTGTATAGCGTGTGCTTGGCGG 4500  
Db |||||  
QY 5090 GTGGGCGGAGGGGCTGTGAGTGTGATGAACCGGCTGTATAGCGTGTGCTTGGCGG 5149  
QY 4501 GGTAAACAGCTCTCCCGCAGCATATGTCTGAGAGCGGCTGACGACGTGTCACT 4560  
Db |||||  
QY 5150 GGTAAACAGCTCTCCCGCAGCATATGTCTGAGAGCGGCTGACGACGTGTCACT 5209  
QY 4561 CAGATCTCTTCTAGTCTTACCATCACTCAGCTCTGAAAGGCTTCCACAGTGGATCAAC 4620  
Db |||||  
QY 5210 CAGATCTCTTCTAGTCTTACCATCACTCAGCTCTGAAAGGCTTCCACAGTGGATCAAC 5269  
QY 4621 GAGGACTGTCCACGCCATGCTCCGCTGTGCTAAGAGATTTTGGATTTGGATATGC 4680  
Db |||||  
QY 5270 GAGGACTGTCCACGCCATGCTCCGCTGTGCTTAAGAGATTTTGGATTTGGATATGC 5329  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAGCTCTGCGCGGATTTGCCGGA 4740

Db |||||  
QY 5330 ACGGTGTTGACTGATTTCAAGGCTGGCTCCAGTCCAGCTCTGCGCGGATTTGCCGGA 5389  
QY 4741 GTCCCTTTCTTCATGTCAACGTGGGTACAAGGGAGTCTGGCGGGCGACGCAATCATG 4800  
Db |||||  
QY 5390 GTCCCTTTCTTCATGTCAACGTGGGTACAAGGGAGTCTGGCGGGCGACGCAATCATG 5449  
QY 4801 CAAACCACTGCTCCCATGTGGAGCAGATCAACGGACATGTGAANAACGGTTCATGAG 4860  
Db |||||  
QY 5450 CAAACCACTGCTCCCATGTGGAGCAGATCAACGGACATGTGAANAACGGTTCATGAG 5509  
QY 4861 ATCTGTTGGGCTTAGACCTGTAGTAACAGTGGCATTGGAAACATTTCCCACTTAACGCGTAC 4920  
Db |||||  
QY 5510 ATCTGTTGGGCTTAGACCTGTAGTAACAGTGGCATTGGAAACATTTCCCACTTAACGCGTAC 5569  
QY 4921 ACCAGGGGCTCTGACGCGCTTCCCGGCGCAAAATTTCTAGGGCGTGTGGCGGGT 4980  
Db |||||  
QY 5570 ACCAGGGGCTCTGACGCGCTTCCCGGCGCAAAATTTCTAGGGCGTGTGGCGGGT 5629  
QY 4981 GCTGCTGAGGATGACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
Db |||||  
QY 5630 GCTGCTGAGGATGACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGGGCATG 5689  
QY 5041 ACCACTGACAAAGTAAAGTGGCGCTGTCAGGTTCCGGCCCCCGAAATTTCTTACAGAAAGT 5100  
Db |||||  
QY 5690 ACCACTGACAAAGTAAAGTGGCGCTGTCAGGTTCCGGCCCCCGAAATTTCTTACAGAAAGT 5749  
QY 5101 GATGGGCTGGTTGACAGGTTACGCTCCAGGTCGCAAAACCCCTCTACGGGAGGAGT 5160  
Db |||||  
QY 5750 GATGGGCTGGTTGACAGGTTACGCTCCAGGTCGCAAAACCCCTCTACGGGAGGAGT 5809  
QY 5161 ACATTTCTGGTGGGCTCAATCAATACCTGTTGGTTCAGCTCCCATGCGAGCCGGA 5220  
Db |||||  
QY 5810 ACATTTCTGGTGGGCTCAATCAATACCTGTTGGTTCAGCTCCCATGCGAGCCGGA 5869  
QY 5221 CCGGAGTACAGTGTCTCACTTCCATGTCAACGACCTCCACATTAACGGCGGAGAG 5280  
Db |||||  
QY 5870 CTGAGCTGAGAGTGTCTCACTTCCATGTCAACGACCTCCACATTAACGGCGGAGAG 5929  
QY 5281 GCTAAGCTGAGCTGGCCAGGGGATCTCCCGCTCTCTGGCCAGTCTCATCAGTACGCCAG 5340  
Db |||||  
QY 5930 GCTAAGCTGAGCTGGCCAGGGGATCTCCCGCTCTCTGGCCAGTCTCATCAGTACGCCAG 5989  
QY 5341 CTGCTGTGGCCTTCTTTGAAGGCAATGCACTACCCGTCAATGATCTCCCGGAGCGTAC 5400  
Db |||||  
QY 5990 CTGCTGTGGCCTTCTTTGAAGGCAATGCACTACCCGTCAATGATCTCCCGGAGCGTAC 6049  
QY 5401 CTCTATGAGGCCAAACCTCTGTGGCGGAGAGTGGCGGGAACATCAACCCGCTGGAG 5460  
Db |||||  
QY 6050 CTCTATGAGGCCAAACCTCTGTGGCGGAGAGTGGCGGGAACATCAACCCGCTGGAG 6109  
QY 5461 TCAGAAATTAAGGTAGTAATTTTGGACTCTTTTCGAGCCGCTCCAGCGGAGGAGTACG 5520  
Db |||||  
QY 6110 TCAGAAATTAAGGTAGTAATTTTGGACTCTTTTCGAGCCGCTCCAGCGGAGGAGTACG 6169  
QY 5521 AGGAAAGTATCGTTCCGGCGGAGATCTCGGAGGTTCAGAAATTTCCCTCGAGCGATG 5580  
Db |||||  
QY 6170 AGGAAAGTATCGTTCCGGCGGAGATCTCGGAGGTTCAGAAATTTCCCTCGAGCGATG 6229  
QY 5581 CCCATATGGGCGACCGCGGATTTACACCTCCACTGTTAGAGTCTCTGGAAGGACCCCGGAC 5640  
Db |||||  
QY 6230 CCCATATGGGCGACCGCGGATTTACACCTCCACTGTTAGAGTCTCTGGAAGGACCCCGGAC 6289  
QY 5641 TAGCTCCCTTCCAGTGTACACGGGTGTCCATTTGCGGCTCTGCAAGGCGCTCCGATACCA 5700  
Db |||||  
QY 6290 TAGCTCCCTTCCAGTGTACACGGGTGTCCATTTGCGGCTCTGCAAGGCGCTCCGATACCA 6349  
QY 5701 CTTCCACGAGGAAAGGAGCGGTGTCTGTGTCAGAAATTTACCGTGTCTTCTGCTTGGCG 5760  
Db |||||  
QY 6350 CTTCCACGAGGAAAGGAGCGGTGTCTGTGTCAGAAATTTACCGTGTCTTCTGCTTGGCG 6409  
QY 5761 GAGCTCGCCACAAAGACCTTCGGAGCTCCGAATCTGCGGCGTCTGACACGCGGACGCA 5820



```
QY 7981 AGATCAAGT 7989
Db 8630 AGATCAAGT 8638

RESULT 11
US-10-029-907-25
; Sequence 25, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-25

Query Match 90.7%; Score 7247.2; DB 4; Length 8638;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 7959; Conservative 0; Mismatches 18; Indels 671; Gaps 3;

QY 2 CCAGCCCCGATTGGGGGCGACATCCACCATAGATCACTCCCTGTGAGGAACACTACTGT 61
Db 2 CCAGCCCCGATTGGGGGCGACATCCACCATAGATCACTCCCTGTGAGGAACACTACTGT 61

QY 62 CTTACGAGAAAGGCTAGCCATGGCGTTAGTAGTGTGTCGAGGCTCCAGGACC 121
Db 62 CTTACGAGAAAGGCTAGCCATGGCGTTAGTAGTGTGTCGAGGCTCCAGGACC 121

QY 122 CCCCTCCCGGAGAGCCATAGTGGTCTCGGAACCGGTGAGTACACCGGAAATGCCAGG 181
Db 122 CCCCTCCCGGAGAGCCATAGTGGTCTCGGAACCGGTGAGTACACCGGAAATGCCAGG 181

QY 182 ACGACCGGTCTCTTCTTGATCAACCCGCTCAATGCCTGGAGATTGGGCGTGCCTCCG 241
Db 182 ACGACCGGTCTCTTCTTGATCAACCCGCTCAATGCCTGGAGATTGGGCGTGCCTCCG 241

QY 242 CGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGTGTAATGCTGTAGGG 301
Db 242 CGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGTGTAATGCTGTAGGG 301

QY 302 TGTCTCGAGTGCCTCGGAGGTCTGTAGACGCTGCACATGAGCAGCAATCCTAAACC 361
Db 302 TGTCTCGAGTGCCTCGGAGGTCTGTAGACGCTGCACATGAGCAGCAATCCTAAACC 361

QY 362 TCAAGAAAAAACCAGGCGGCCATGATTGAACAGATGGAATTCACGCGGTTCTCC 421
Db 362 TCAAGAAAAAACCAGGCGGCCATGATTGAACAGATGGAATTCACGCGGTTCTCC 421

QY 422 GCGCGTGGTGGAGAGGCTATTCGGCTATGATCGGGCAACAGACAAATCGGCTGCTC 481
Db 422 GCGCGTGGTGGAGAGGCTATTCGGCTATGATCGGGCAACAGACAAATCGGCTGCTC 481

QY 482 TGATGCCCGCTTCCGGTGTGAGCGAGGCGCGCGGTTCTTTTGTCAAGACCGA 541
Db 482 TGATGCCCGCTTCCGGTGTGAGCGAGGCGCGCGGTTCTTTTGTCAAGACCGA 541

QY 542 CTTGTCCGCTGCCCTGAATGAATCTGAGGACGAGGCGCGGCTATCGTGGCTGCCAC 601
Db 542 CTTGTCCGCTGCCCTGAATGAATCTGAGGACGAGGCGCGGCTATCGTGGCTGCCAC 601

QY 602 GACGGCGGTTCCCTTGGCAGCTGTCTCAAGCTTGTCACTGAAGCGGGAAGGACCTGGCT 661
Db 602 GACGGCGGTTCCCTTGGCAGCTGTCTCAAGCTTGTCACTGAAGCGGGAAGGACCTGGCT 661

QY 662 GCTATTGGCGGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTGCCAGAA 721
Db 662 GCTATTGGCGGAAGTCCCGGGCAGGATCTCTGTCTCATCTCACCTTGTCTCTGCCAGAA 721

QY 722 AGTATCCATCATGGCTGATGAAATCGCGGCTGTGATACGCTTGTATCCGGCTACCTGCC 781
Db 722 AGTATCCATCATGGCTGATGAAATCGCGGCTGTGATACGCTTGTATCCGGCTACCTGCC 781

QY 782 ATTGACACCAAGCGAAACATCGCATCGAGGACGATCTCGATCGGATGGAAGCGGCTCT 841
Db 782 ATTGACACCAAGCGAAACATCGCATCGAGGACGATCTCGATCGGATGGAAGCGGCTCT 841

QY 842 TGTTCGATCAGGATGATCTGGAAGGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCG 901
Db 842 TGTTCGATCAGGATGATCTGGAAGGAGCATCAGGGGCTCGCGCCAGCCGAACTGTTCG 901

QY 902 CAGGCTCAAGGCGGATGCCCGAGCGGAGGATCTCGTCTGACCCATGGCGATGCTG 961
Db 902 CAGGCTCAAGGCGGATGCCCGAGCGGAGGATCTCGTCTGACCCATGGCGATGCTG 961

QY 962 CTTGCGGATATCATGGTGGAAAAATGGCGCTTTCTGATTCATCATCTGTGGCGGCT 1021
Db 962 CTTGCGGATATCATGGTGGAAAAATGGCGCTTTCTGATTCATCATCTGTGGCGGCT 1021

QY 1022 GGGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATGCTGAAGAGCT 1081
Db 1022 GGGTGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCCGCTGATATGCTGAAGAGCT 1081

QY 1082 TGGCGCGAATGGGTGACCGCTTCTCTGTCTTACGATTCGCGCTCCCGATTCGCA 1141
Db 1082 TGGCGCGAATGGGTGACCGCTTCTCTGTCTTACGATTCGCGCTCCCGATTCGCA 1141

QY 1142 GGCATCGCTTCTATCGGCTTCTTGAAGAGTCTCTCTGAGTT-----TAAAC 1189
Db 1142 GGCATCGCTTCTATCGGCTTCTTGAAGAGTCTCTCTGAGTT-----TAAAC 1189

QY 1190 AGACACAAACGGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCGCTCCCGCCCTTA 1249
Db 1202 AGACACAAACGGTTTCCCTCTAGCGGATCAATTCG-----CCCCCCCCCTTA 1250

QY 1250 ACGTTACTGGCGGAAAGCGCTTGGAAATAGGCGCGGTGCGTTGTCTATATGTTATTT 1309
Db 1251 ACGTTACTGGCGGAAAGCGCTTGGAAATAGGCGCGGTGCGTTGTCTATATGTTATTT 1310

QY 1310 CCACCATATTTGGCGTCTTTTGGCAATGTAGGCGCGGAAACCTGCGCTGCTCTTCTGA 1369
Db 1311 CCACCATATTTGGCGTCTTTTGGCAATGTAGGCGCGGAAACCTGCGCTGCTCTTCTGA 1370

QY 1370 CGAGCATTTCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCG 1429
Db 1371 CGAGCATTTCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCG 1430

QY 1430 TGAAGAGACAGTTCTCTGGAAGCTTTTGAAGACAAACACGCTGTAGGACCGCTTT 1489
Db 1431 TGAAGAGACAGTTCTCTGGAAGCTTTTGAAGACAAACACGCTGTAGGACCGCTTT 1490

QY 1490 GCAGGACGGAACCCCCCAGCTGGGACAGGTGCTCTGCGGCCAAAGACCGCTGAT 1549
Db 1491 GCAGGACGGAACCCCCCAGCTGGGACAGGTGCTCTGCGGCCAAAGACCGCTGAT 1550

QY 1550 AAGATACACCTCAAGGCGGACAAACCCAGTGGCCAGTCTGAGTTGATGTTGTCG 1609
Db 1551 AAGATACACCTCAAGGCGGACAAACCCAGTGGCCAGTCTGAGTTGATGTTGTCG 1610

QY 1610 AAAGAGTCAAAATGGCTCTCTCAAGCTATTCAAAGGGGCTGAAGGATGCCCAAGG 1669
Db 1611 AAAGAGTCAAAATGGCTCTCTCAAGCTATTCAAAGGGGCTGAAGGATGCCCAAGG 1670

QY 1670 TACCCCAATTGATGGGATCTGATCTGGGGCTCGGTGACATGCTTTTACATGTTTGTAGT 1729
```



QY 3242 AACGCCCTCGGCGATGTTCCGATTCCTCGTCTCTGTCGAGTCTATGACGCGGGCTGTG 3301  
Db AACGGCCCTCGGCGATGTTCCGATTCCTCGTCTCTGTCGAGTCTATGACGCGGGCTGTG 3950  
QY 3302 CTTGGTACGAGCTACAGCCCGCCGAGACCTCACTAGGTTGCGGGCTTACCTAAACACAC 3361  
Db CTTGGTACGAGCTACAGCCCGCCGAGACCTCACTAGGTTGCGGGCTTACCTAAACACAC 4010  
QY 3362 CAGGGTTGCGGCTCTGCCAGGACCACTGGAGTTCTGGGAGCGCTCTTTACAGGCTCA 3421  
Db CAGGGTTGCGGCTCTGCCAGGACCACTGGAGTTCTGGGAGCGCTCTTTACAGGCTCA 4070  
QY 3422 CCCACATAGACGCCCATTTCTGTCGACACTAAGCAGCAGCAGACCACTTCCCTACC 3481  
Db CCCACATAGACGCCCATTTCTGTCGACACTAAGCAGCAGCAGACCACTTCCCTACC 4130  
QY 3482 TGGTAGCATACCAAGGCTACGGTGTGGGCCAGGGCTCAGGCTCCACCTCCATCGTGGGACC 3541  
Db TGGTAGCATACCAAGGCTACGGTGTGGGCCAGGGCTCAGGCTCCACCTCCATCGTGGGACC 4190  
QY 3542 AAATGTGGAAGTGTCTATACGGCTAAAGCTAAGCTGACGCGGCCAAGCCCTCTGT 3601  
Db AAATGTGGAAGTGTCTATACGGCTAAAGCTAAGCTGACGCGGCCAAGCCCTCTGT 4250  
QY 3602 ATAGGCTGGAGCCGTTCAAACGAGTTACTACACACACCCCATAAACCAATATATCA 3661  
Db ATAGGCTGGAGCCGTTCAAACGAGTTACTACACACACCCCATAAACCAATATATCA 4310  
QY 3662 TGGCATGCAATGTCGGCTGACCTGGAGTCTGACGAGCACCTGGGTGCTGGTAGGGGAG 3721  
Db TGGCATGCAATGTCGGCTGACCTGGAGTCTGACGAGCACCTGGGTGCTGGTAGGGGAG 4370  
QY 3722 TCTTAGAGCTCTGGCCGCTATTGCTGTGACAAACAGCAGCGTGTGATTTGGGAGGA 3781  
Db TCTTAGAGCTCTGGCCGCTATTGCTGTGACAAACAGCAGCGTGTGATTTGGGAGGA 4430  
QY 3782 TCATCTTTGCGGAAGCCGCGCATCATCTCCGACAGGGAAGTCTTTACCGGAGTTCG 3841  
Db TCATCTTTGCGGAAGCCGCGCATCATCTCCGACAGGGAAGTCTTTACCGGAGTTCG 4490  
QY 3842 ATGAGATGGAAGTGTGCGCTCAACCTCCCTTACATCGAAACAGGAATCGACTCGCCG 3901  
Db ATGAGATGGAAGTGTGCGCTCAACCTCCCTTACATCGAAACAGGAATCGACTCGCCG 4550  
QY 3902 AACAAATCAAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAGCGAGGCTG 3961  
Db AACAAATCAAACAGAGGCAATCGGTTGCTGCAAAACAGCCACCAAGCAGCGAGGCTG 4610  
QY 3962 CTGCTCCGTTGTTGAATCCAGTGGGAGCCCTCGAAGCTTCTGGGCGAAGCATATGT 4021  
Db CTGCTCCGTTGTTGAATCCAGTGGGAGCCCTCGAAGCTTCTGGGCGAAGCATATGT 4670  
QY 4022 GGAATTTTCATCAGCGGATACAAATATTTAGCAGGCTTGTCCACTGCTGCTGGCAACCCCG 4081  
Db GGAATTTTCATCAGCGGATACAAATATTTAGCAGGCTTGTCCACTGCTGCTGGCAACCCCG 4730  
QY 4082 CGATAGCATCATGTAGGATTCACAGCTCTATCACAGCCCGCTCACACCACCAATATA 4141  
Db CGATAGCATCATGTAGGATTCACAGCTCTATCACAGCCCGCTCACACCACCAATATA 4790  
QY 4142 CCCTCCCTGTTTAACTCTGGGGGATGGGTGGCGGCCCAACTGCTCTCCAGGCTG 4201  
Db CCCTCCCTGTTTAACTCTGGGGGATGGGTGGCGGCCCAACTGCTCTCCAGGCTG 4850  
QY 4202 CTTCTGCTTTCTGAGGCGCCGCTCGCTGGAGCGGCTGTGGCAGCATAGGCTTTGGGA 4261  
Db CTTCTGCTTTCTGAGGCGCCGCTCGCTGGAGCGGCTGTGGCAGCATAGGCTTTGGGA 4910  
QY 4262 AGTGCTGTGGATATTTTGGCAGGTTATGGACAGGGGTGGCAGGCGGCTCTGCGCT 4321  
Db AGTGCTGTGGATATTTTGGCAGGTTATGGACAGGGGTGGCAGGCGGCTCTGCGCT 4970

QY 4322 TTAAAGTATGATGCGCGAGATGSCCTCCACAGGACCTGGTTAACTACTTCCCTGCTA 4381  
Db TTAAAGTATGATGCGCGAGATGSCCTCCACAGGACCTGGTTAACTACTTCCCTGCTA 5030  
QY 4382 TCTCTCTCCCTCGGCGCTAGTCTCGGGTCTGTGCGCAGCAGATCTGCGTCGCGCAGC 4441  
Db TCTCTCTCCCTCGGCGCTAGTCTCGGGTCTGTGCGCAGCAGATCTGCGTCGCGCAGC 5090  
QY 4442 TGGGCCCAAGGAGGGGGCTGTGAGTGATGAACCGGCTGATAGGCTTTCGCTTCGCGG 4501  
Db TGGGCCCAAGGAGGGGGCTGTGAGTGATGAACCGGCTGATAGGCTTTCGCTTCGCGG 5150  
QY 4502 GTAACCACTCTCCCCCAGCAGCTATGCTGAGAGCAGCAGCTGAGCAGCTGCTCACTC 4561  
Db GTAACCACTCTCTCCCCCAGCAGCTATGCTGAGAGCAGCAGCTGAGCAGCTGCTCACTC 5210  
QY 4562 AGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTTCACAGTGGATCAACG 4621  
Db AGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTTCACAGTGGATCAACG 5270  
QY 4622 AGGACTGCTCAGCCCATGCTCGGCTCGTGGCTAAGAGATGTTTGGGATTTGATATGCA 4681  
Db AGGACTGCTCAGCCCATGCTCGGCTCGTGGCTAAGAGATGTTTGGGATTTGATATGCA 5330  
QY 4682 CGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCCGCGGATTCGCCGGAG 4741  
Db CGGTGTTGACTGATTTCAAGAGCTGCTCCAGTCCAGTCCCGCGGATTCGCCGGAG 5390  
QY 4742 TCCCTCTTCTCATGTCAACGTGGGTACAAGGAGTCTTGGCGGGGCGACGGCATCATGC 4801  
Db TCCCTCTTCTCATGTCAACGTGGGTACAAGGAGTCTTGGCGGGGCGACGGCATCATGC 5450  
QY 4802 AAACACCTGCCCATGTGGAGACAGATCAACGGACATGTGAAGACGGTTCCATGAGGA 4861  
Db AAACACCTGCCCATGTGGAGACAGATCAACGGACATGTGAAGACGGTTCCATGAGGA 5510  
QY 4862 TGTGGGGCTTAGGACTGTAGTAAACAGTGGCATGGAACATTTCCCATTAACCGGTACA 4921  
Db TGTGGGGCTTAGGACTGTAGTAAACAGTGGCATGGAACATTTCCCATTAACCGGTACA 5570  
QY 4922 CCACGGGCCCTCTGACCGCCCTCCCGCGGCCAAATATTTCTAGGGCGCTGTGGCGGTGG 4981  
Db CCACGGGCCCTCTGACCGCCCTCCCGCGGCCAAATATTTCTAGGGCGCTGTGGCGGTGG 5630  
QY 4982 CTGCTGAGGATGCTGAGGTTACGCGGGTGGGGATTTCCACTAGTGAACGGGATGA 5041  
Db CTGCTGAGGATGCTGAGGTTACGCGGGTGGGGATTTCCACTAGTGAACGGGATGA 5690  
QY 5042 CCACGACAAAGTAAAGTCCCGTGTACAGTTCCGCGCCCGGAAATTTCTACAGAAAGTGG 5101  
Db CCACGACAAAGTAAAGTCCCGTGTACAGTTCCGCGCCCGGAAATTTCTACAGAAAGTGG 5750  
QY 5102 ATGGGTGCGGTTGACAGGTACGCTCCAGCGGTGCAAAACCCCTCTACGCGGAGGATCA 5161  
Db ATGGGTGCGGTTGACAGGTACGCTCCAGCGGTGCAAAACCCCTCTACGCGGAGGATCA 5810  
QY 5162 CATTCCTGTCGGGCTCAATCAATACCTGTTGGGTCAAGTCCCATGCGAGCCCGAAC 5221  
Db CATTCCTGTCGGGCTCAATCAATACCTGTTGGGTCAAGTCCCATGCGAGCCCGAAC 5870  
QY 5222 CGGACGTAGCAGTGTCTTCCATGCTCAGCGCCCTCCACATTTACGGCGGACCG 5281  
Db CGGACGTAGCAGTGTCTTCCATGCTCAGCGCCCTCCACATTTACGGCGGACCG 5930  
QY 5282 CTAAGCGTAGGCTGGCCAGGGATCTTCCCTCTTGGTGGCAGCTCATCAGCTAGCCAGC 5341  
Db CTAAGCGTAGGCTGGCCAGGGATCTTCCCTCTTGGTGGCAGCTCATCAGCTAGCCAGC 5990  
QY 5342 TGTCTGCGCTTCTTGAAGGCAATGACATACCGTCAATGACTCCCGGAGCGCTGACC 5401  
Db TGTCTGCGCTTCTTGAAGGCAATGACATACCGTCAATGACTCCCGGAGCGCTGACC 6050  
QY 5402 TCATCGAGGCCAACCTCTCTGTGGCGGACAGGATGGCGGGGAACATCACCCGCTGAGT 5461





```
QY 7622 TATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAAGCCTGTCTCGTG 7681
Db 8271 TATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAAGCCTGTCTCGTG 8330
QY 7682 CCGACCCCGCTGGTTCATGTGGTGGCTACTCCCTACTTTCTGTAGGGTAGGCATCTATC 7741
Db 8331 CCGACCCCGCTGGTTCATGTGGTGGCTACTCCCTACTTTCTGTAGGGTAGGCATCTATC 8390
QY 7742 TATCCCCCAACCGATGAACCGGGACCTTAAACACTCCAGGCCAATPAGGCCATCTCTGTTTT 7801
Db 8391 TACTCCCCCAACCGATGAACCGGGAGCTTAAACACTCCAGGCCAATPAGGCCATCTCTGTTTT 8450
QY 7802 TTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7861
Db 8451 TTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 8510
QY 7862 TTTTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7921
Db 8511 TTTTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 8570
QY 7922 AGTGTGAAGGTCGGTAGCCGCTTGACTGCAAGAGAGTGTGATAGTGGCCTCTCTGCA 7981
Db 8571 AGTGTGAAGGTCGGTAGCCGCTTGACTGCAAGAGAGTGTGATAGTGGCCTCTCTGCA 8630
QY 7982 GATCAAGT 7989
Db 8631 GATCAAGT 8638

RESULT 12
US-10-029-907-2
; Sequence 2, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; TITLE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 8642
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
; NAME/KEY: variation
; LOCATION: 6268
; OTHER INFORMATION: r = a or g
; NAME/KEY: variation
; LOCATION: 4446
; OTHER INFORMATION: r = a or g
US-10-029-907-2

Query Match 90.78; Score 7245.2; DB 4; Length 8642;
Best Local Similarity 92.18; Pred. No. 0;
Matches 7965; Conservative 2; Mismatches 10; Indels 675; Gaps 4;

QY 2 CCAGCCCCCGATGGGGGGGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTGT 61
Db 2 CCAGCCCCCGATGGGGGGGACACTCCACCATAGATCACTCCCTGTGAGGAACTACTGT 61
QY 62 CTTACCGCAAGAGCGTCTAGCCATGCGGTAGTATGAGTGTGAGCCTCCAGGACC 121
Db 62 CTTACCGCAAGAGCGTCTAGCCATGCGGTAGTATGAGTGTGAGCCTCCAGGACC 121
QY 122 CCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGG 181
```

```
Db 122 CCCCTCCCGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGGAATTGCCAGG 181
QY 182 ACGACCGGGTCCCTTTCTTTGGATCAACCCGCTCAATGCTGAGATTTGGGCGTCCCGCCG 241
Db 182 ACGACCGGGTCCCTTTCTTTGGATCAACCCGCTCAATGCTGAGATTTGGGCGTCCCGCCG 241
QY 242 CGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGGTACTGCTGCTGATAGGG 301
Db 242 CGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGGCTTGTGGTACTGCTGCTGATAGGG 301
QY 302 TGTTCGAGTGCCTCCGGGAGGTCTCGTAGCCGTGCACCATGAGCAGCAATCTTAACC 361
Db 302 TGTTCGAGTGCCTCCGGGAGGTCTCGTAGCCGTGCACCATGAGCAGCAATCTTAACC 361
QY 362 TCAAGAAAAACAAAGGGCGCGCATGATTGAACAGATGGATTGACACGAGGTTCTCC 421
Db 362 TCAAGAAAAACAAAGGGCGCGCATGATTGAACAGATGGATTGACACGAGGTTCTCC 421
QY 422 GCGCGCTTGGTGGAGAGGCTATTGGGCTATGACTGGGCAACAAGCAATCGGCTGCTC 481
Db 422 GCGCGCTTGGTGGAGAGGCTATTGGGCTATGACTGGGCAACAAGCAATCGGCTGCTC 481
QY 482 TGATGCGCGCTGTTCGCGCTGTGAGCGAGGGGCGCGCTTCTTTTGTCAAGACCGA 541
Db 482 TGATGCGCGCTGTTCGCGCTGTGAGCGAGGGGCGCGCTTCTTTTGTCAAGACCGA 541
QY 542 CCTGTCCGCTGCCCTGAATGAATGCAGAGCAGGCGCGGTATCGTGGTGGCCAC 601
Db 542 CCTGTCCGCTGCCCTGAATGAATGCAGAGCAGGCGCGGTATCGTGGTGGCCAC 601
QY 602 GACGGGCTTCTTGGCAGCTGTGTCGACGTTGTCACGAGCGGGAAGGACTGGCT 661
Db 602 GACGGGCTTCTTGGCAGCTGTGTCGACGTTGTCACGAGCGGGAAGGACTGGCT 661
QY 662 GCTATTGGGCGAAGTCCGGGCGAGGATCTCTGTGTCATCTCAGCTTCTCCTCCGAGAA 721
Db 662 GCTATTGGGCGAAGTCCGGGCGAGGATCTCTGTGTCATCTCAGCTTCTCCTCCGAGAA 721
QY 722 AGTATCCATCATGGCTGATGCAATGGCGGCTGTCATACGCTTGTGATCCGCTACCTGCC 781
Db 722 AGTATCCATCATGGCTGATGCAATGGCGGCTGTCATACGCTTGTGATCCGCTACCTGCC 781
QY 782 ATTTCGACCAACGAGGAAACATCGATCGAGCGAGCACGCTACTCGATGGAAGCGGCT 841
Db 782 ATTTCGACCAACGAGGAAACATCGATCGAGCGAGCACGCTACTCGATGGAAGCGGCT 841
QY 842 TGTTCGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTCCG 901
Db 842 TGTTCGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTCCG 901
QY 902 CAGGCTCAAGGCGCGCATGCCGAGCAGGATCTGCTGTCGATACCCATGCGCATGCTG 961
Db 902 CAGGCTCAAGGCGCGCATGCCGAGCAGGATCTGCTGTCGATACCCATGCGCATGCTG 961
QY 962 CTTGCGCAATATCATGGTGGAAAATGGCCCTTTCTGAGTTTCTGAGTTTCTGAGTT 1021
Db 962 CTTGCGCAATATCATGGTGGAAAATGGCCCTTTCTGAGTTTCTGAGTTTCTGAGTT 1021
QY 1022 GGGTGTGGGCGCGCTATCAGGACATAGGTTGGCTACCCGTGATATGCTGCAAGAGCT 1081
Db 1022 GGGTGTGGGCGCGCTATCAGGACATAGGTTGGCTACCCGTGATATGCTGCAAGAGCT 1081
QY 1082 TGGCGCGAATGGGCTGACCGCTTCTCGTGTCTTTTACGGTATCGCGCTCCCGATTGCA 1141
Db 1082 TGGCGCGAATGGGCTGACCGCTTCTCGTGTCTTTTACGGTATCGCGCTCCCGATTGCA 1141
QY 1142 GCGCATGCGCTTCTATCGCTTCTTACGAGTCTTCTGAGTTTCTGAGTTTCTGAGTT 1189
Db 1142 GCGCATGCGCTTCTATCGCTTCTTACGAGTCTTCTGAGTTTCTGAGTTTCTGAGTT 1201
QY 1190 AGACCAACAGGTTTCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCGCCCTTA 1249
```

Db 1202 AGACACAACGGTTTCCCTCTAGCGGATCAATTCCG-----CCCCCCCCCTTA 1250  
QY 1250 ACGTTACTGGCCGAAGCCGCTGGATAGAGCGCGTGTGCGTTTGTCTATATATATTTT 1309  
Db 1251 ACGTTACTGGCCGAAGCCGCTGGATAGAGCGCGTGTGCGTTTGTCTATATATATTTT 1310  
QY 1310 CCACCATATTCGCGTCTTTTGGCAATGTAGGCGCCGGAACCTGGCCCTGTCTTTGA 1369  
Db 1311 CCACCATATTCGCGTCTTTTGGCAATGTAGGCGCCGGAACCTGGCCCTGTCTTTGA 1370  
QY 1370 CGAGCATTCCTAGGGTCTTTTCCCTCTCGCAAGAGAAAGTCAAGTCTGTGAATGCG 1429  
Db 1371 CGAGCATTCCTAGGGTCTTTTCCCTCTCGCAAGAGAAAGTCAAGTCTGTGAATGCG 1430  
QY 1430 TGAAGAGACAGTCTCTCGAAGCTTCTTGAAGCAACACAGCTGTAGCGACCTTT 1489  
Db 1431 TGAAGAGACAGTCTCTCGAAGCTTCTTGAAGCAACACAGCTGTAGCGACCTTT 1490  
QY 1490 GCAGGACGGAACCCCACTGGCGACAGGTGCTCTGCGGCCAAAGACCACTGTAT 1549  
Db 1491 GCAGGACGGAACCCCACTGGCGACAGGTGCTCTGCGGCCAAAGACCACTGTAT 1550  
QY 1550 AAGATACCTCGAAGCGGCAACCCCACTGGCGACAGGTGCTCTGCGGCCAAAGACCACTGTAT 1609  
Db 1551 AAGATACCTCGAAGCGGCAACCCCACTGGCGACAGGTGCTCTGCGGCCAAAGACCACTGTAT 1610  
QY 1610 AAGAGTCAAAAGCTCTCTCAAGCTATTCAACAGGGCTGAAGATGCCAGAGG 1669  
Db 1611 AAGAGTCAAAAGCTCTCTCAAGCTATTCAACAGGGCTGAAGATGCCAGAGG 1670  
QY 1670 TACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGTTTGT 1729  
Db 1671 TACCCCATTTGATGGGATCTGATCTGGGCTCGGTGCACATGCTTTACATGTTTGT 1730  
QY 1730 CGAGGTTAAAGAGTCTAGGCCCCCGGAAACCAAGGAGCGTGGTTTCTTTGAAAAAC 1789  
Db 1731 CGAGGTTAAAGAGTCTAGGCCCCCGGAAACCAAGGAGCGTGGTTTCTTTGAAAAAC 1790  
QY 1790 ACGATAATACCATG----- 1803  
Db 1791 ACGATAATACCATGAGCCGGAGATGGAGCATCGTGGGAGCGCGGTTCGTAGGTC 1850  
QY 1804 ----- 1803  
Db 1851 TGATATCTTTGACCTTGTACCGCACTATAAGCTGTTCTCGTAGGCTCATATGTTGT 1910  
QY 1804 ----- 1803  
Db 1911 TACAATATTTATCACCAGGCGGAGGCACACTTGCAAGTGTGGATCCCGCCCTCAAG 1970  
QY 1804 ----- 1803  
Db 1971 TTGGGGGGCGCGATGCGGTATCTCTCCTCAGTGGCGGATCCACCAGAGCTAATCT 2030  
QY 1804 ----- 1803  
Db 2031 TTACCANTACCAAAATCTTTGCGCATACTCGGTCCACTCATGTGTCTCAGGCTGTA 2090  
QY 1804 ----- 1803  
Db 2091 TAACCAAGTCCGTACTTCTGTCGGCGACACGGGCTCATTCGTGCATGCATGCTGGTC 2150  
QY 1804 ----- 1803  
Db 2151 GGAAGTTGCTGGGGTCAATTATGTCCAAATGGCTCTCATGAAGTTGGCGGCACTGACAG 2210  
QY 1804 ----- 1803  
Db 2211 GTACGTACGTTTATGACCATCTCAOCCCACTGCGGACTGGGCCACGCGGCTACGAG 2270  
QY 1804 ----- 1803  
Db 2271 ACCTTGGGTGGAGTTGAGCCCGTGTCTTCTCTGATATGAGACCAAGGTTATCACCT 2330

QY 1804 ----- 1803  
Db 2331 GGGGGGAGACACACCGCGGTGTGGGACATCATCTTGGGCTGCCGTCTCGGCCGCA 2390  
QY 1804 ----- 1803  
Db 2391 GGGGAGGAGATACATCTGGGACCGGAGACAGCCTTTGAAGGGCAGGGGTGGACATCC 2450  
QY 1804 --GCGCCTTATACGCGCTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCTA 1861  
Db 2451 TCGCGCCTTATACGCGCTACTCCCAACAGACGCGAGGCTACTTGGCTGCATCATCTA 2510  
QY 1862 GCTCACAAGCGCGGACAGGAACAGGTCTGAGGGAGGTCCAAAGTGGTCTCCACGCAA 1921  
Db 2511 GCTCACAAGCGCGGACAGGAACAGGTCTGAGGGAGGTCCAAAGTGGTCTCCACGCAA 2570  
QY 1922 CACAATCTTCTGCGACCTGCGTCAATGGCGTGTGTGGACTGTCTATCATGTGCGG 1981  
Db 2571 CACAATCTTCTGCGACCTGCGTCAATGGCGTGTGTGGACTGTCTATCATGTGCGG 2630  
QY 1982 GCTCAAGACCTTTCGCGGCCCAAGGGCCCAATCACCACCAATGTACACCAATGTGGAC 2041  
Db 2631 GCTCAAGACCTTTCGCGGCCCAAGGGCCCAATCACCACCAATGTACACCAATGTGGAC 2690  
QY 2042 AGGACTCTGTCGCTGGCAAGCGCCCCCGGGCGGTTCCTTGACACCATGCACCTCGG 2101  
Db 2691 AGGACTCTGTCGCTGGCAAGCGCCCCCGGGCGGTTCCTTGACACCATGCACCTCGG 2750  
QY 2102 GCAGCTCGGACCTTTACTTTGTGTCACGAGCATGCCATGTCTCATTCGGTGGCGCGCGG 2161  
Db 2751 GCAGCTCGGACCTTTACTTTGTGTCACGAGCATGCCATGTCTCATTCGGTGGCGCGCGG 2810  
QY 2162 GCGACAGCAGGGGAGCTTCTCTCCCGAGGCGCTCTCCCTACTTGAAGGCTCTTCGG 2221  
Db 2811 GCGACAGCAGGGGAGCTTCTCTCCCGAGGCGCTCTCCCTACTTGAAGGCTCTTCGG 2870  
QY 2222 GGGTCCACTGTCTGTCGCCCTCGGGCACGCTGTGGGCACTTTTCGGGTGCGCGTGTGA 2281  
Db 2871 GGGTCCACTGTCTGTCGCCCTCGGGCACGCTGTGGGCACTTTTCGGGTGCGCGTGTGA 2930  
QY 2282 CCGGAGGGTTGCGAAGCGGTGGACTTTGTACCCGTCAGTCTATGGAACCACTATGC 2341  
Db 2931 CCGGAGGGTTGCGAAGCGGTGGACTTTGTACCCGTCAGTCTATGGAACCACTATGC 2990  
QY 2342 GGTCCCGGTCTTTCAGGACACTCGTCCCTCGCGCGGTACCGGAGACTTCCAGGTGC 2401  
Db 2991 GGTCCCGGTCTTTCAGGACAACTCGTCCCTCGCGCGGTACCGGAGACTTCCAGGTGC 3050  
QY 2402 CCGATCTACCGCCCTACTGCTAGCGGCAAGAGCACTAAGGTGCGGCTGGGTATGCAG 2461  
Db 3051 CCGATCTACCGCCCTACTGCTAGCGGCAAGAGCACTAAGGTGCGGCTGGGTATGCAG 3110  
QY 2462 CCGAAGGGTATAAGGTGTGCTGTGAACCCGTCGTCGCCGCACTTAGGTTCGGGG 2521  
Db 3111 CCGAAGGGTATAAGGTGTGCTGTGAACCCGTCGTCGCCGCACTTAGGTTCGGGG 3170  
QY 2522 CGTATATCTTAAGGACATGATGTATCGACCTTAAATCAGAACCGGGTAAAGACATCA 2581  
Db 3171 CGTATATCTTAAGGACATGATGTATCGACCTTAAATCAGAACCGGGTAAAGACATCA 3230  
QY 2582 CCACGGGTGCCCCCATCTACGTACTCCACTATGGCAAGTTTCTTCCGACGCTGGTGTCT 2641  
Db 3231 CCACGGGTGCCCCCATCTACGTACTCCACTATGGCAAGTTTCTTCCGACGCTGGTGTCT 3290  
QY 2642 CTGGGGGCGCTTATGACATCAATAATGTATGTATGTCACCTCACTGATGCTGACACATA 2701  
Db 3291 CTGGGGGCGCTTATGACATCAATAATGTATGTATGTCACCTCACTGATGCTGACACATA 3350  
QY 2702 TCCTGGGATCGGACAGTCTCTGGACCAAGCGGAGACGCTGGAGCGGACTCGTGTGC 2761  
Db 3351 TCCTGGGATCGGACAGTCTCTGGACCAAGCGGAGACGCTGGAGCGGACTCGTGTGC 3410

QY 2762 TCGCACCGCTACGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGAGGTGG 2821  
Db 3411 TCGCCACCGCTACGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGAGGTGG 3470  
QY 2822 CTCGTCCAGCACTGGAGAAATCCCTTTTATGGCAAAAGCCATCCCATCGAGACCATCA 2881  
Db 3471 CTCGTCCAGCACTGGAGAAATCCCTTTTATGGCAAAAGCCATCCCATCGAGACCATCA 3530  
QY 2882 AGGGGGGAGGACCTCATTTTCTGCCATTCGAAGAAATGTGATGAGCTCGCGCGGA 2941  
Db 3531 AGGGGGGAGGACCTCATTTTCTGCCATTCGAAGAAATGTGATGAGCTCGCGCGGA 3590  
QY 2942 AGCTGTCGGCTCGGACTCAATGCTGTAGCATATTAACCGGGCTTGATGATCGGTCA 3001  
Db 3591 AGCTGTCGGCTCGGACTCAATGCTGTAGCATATTAACCGGGCTTGATGATCGGTCA 3650  
QY 3002 TACCAACTAGCGGAGAGCTCATTTGCTGTAGCAACGACGCTCTAAAGACGGGCTTTACCG 3061  
Db 3651 TACCAACTAGCGGAGAGCTCATTTGCTGTAGCAACGACGCTCTAAAGACGGGCTTTACCG 3710  
QY 3062 GCGATTTGCACTGATGATCGATGCAATACATGTGTACCCAGACAGTTCAGCC 3121  
Db 3711 GCGATTTGCACTGATGATGCAATACATGTGTACCCAGACAGTTCAGCC 3770  
QY 3122 TGGACCGGACCTTACCATTTAGACGACGACCGTGCACAGACGCGGTGTACGCTCGC 3181  
Db 3771 TGGACCGGACCTTACCATTTAGACGACGACCGTGCACAGACGCGGTGTACGCTCGC 3830  
QY 3182 AGCGGGCAGGACGACTGGTAGGGCAGGATGGGCATTTACAGGTTTGTGACTCCAGGAG 3241  
Db 3831 AGCGGGCAGGACGACTGGTAGGGCAGGATGGGCATTTACAGGTTTGTGACTCCAGGAG 3890  
QY 3242 AACGGCCCTCGGGATGTTGCAATTCCTCGGTTCTGTGCGAGTGCTATGACGCGGCTGTG 3301  
Db 3891 AACGGCCCTCGGGATGTTGCAATTCCTCGGTTCTGTGCGAGTGCTATGACGCGGCTGTG 3350  
QY 3302 CTTGGTACGAGCTACGCGCGCGGACCTCAGTTAGGTTGCGGCTTACTTAAACACAC 3361  
Db 3951 CTTGGTACGAGCTACGCGCGCGGACCTCAGTTAGGTTGCGGCTTACTTAAACACAC 4010  
QY 3362 CAGGGTTGCCGCTCTGCGAGGACCATCTGGAGTTCTGGGAGAGCGCTTTTACAGGCTCA 3421  
Db 4011 CAGGGTTGCCGCTCTGCGAGGACCATCTGGAGTTCTGGGAGAGCGCTTTTACAGGCTCA 4070  
QY 3422 CCCACATAGACGCCATTTCTTGTCCAGACTAAGCAGGACGAGACAACTTCCCTPACC 3481  
Db 4071 CCCACATAGACGCCATTTCTTGTCCAGACTAAGCAGGACGAGACAACTTCCCTPACC 4130  
QY 3482 TGGTAGCATACGAGCTACGCTGTGCGCAGGCTCAGGCTCCACCTCCATCGTGGGACC 3541  
Db 4131 TGGTAGCATACGAGCTACGCTGTGCGCAGGCTCAGGCTCCACCTCCATCGTGGGACC 4190  
QY 3542 AAATGTGGAAGTGTCTCATACGCTAAAGCTTACGCTGCACGGGCCAACGCCCTGCTGT 3601  
Db 4191 AAATGTGGAAGTGTCTCATACGCTAAAGCTTACGCTGCACGGGCCAACGCCCTGCTGT 4250  
QY 3602 ATAGGCTGGAGCGGTTCAAAAAGAGTTACTACCAACACCCCATTAACCAATATACATCA 3661  
Db 4251 ATAGGCTGGAGCGGTTCAAAAAGAGTTACTACCAACACCCCATTAACCAATATACATCA 4310  
QY 3662 TGGCATCATGTGCGCTGACTGGAGTGTGTCAGGACCTGAGGACCTGAGTGGCGGAG 3721  
Db 4311 TGGCATCATGTGCGCTGACTGGAGTGTGTCAGGACCTGAGGACCTGAGTGGCGGAG 4370  
QY 3722 TCCTAGCAGCTCTGCGCGGCTATTGCTGTGCAACAGGACGCTGGTCAATTTGGGCGAGGA 3781  
Db 4371 TCCTAGCAGCTCTGCGCGGCTATTGCTGTGCAACAGGACGCTGGTCAATTTGGGCGAGGA 4430  
QY 3782 TCACTTTGTCCGGAAGCCGCGCATCAATCCCGACAGGAGTCTTTACCGGAGTTTCG 3841  
Db 4431 TCACTTTGTCCGGAAGCCGCGCATCAATCCCGACAGGAGTCTTTACCGGAGTTTCG 4490  
QY 3842 ATGAGATGGAGAGTGGCGCTCACCTCCCTTACATCGAAACAGGGAATGACGCTCGCGG 3901

Db 4491 ATGAGATGGAGAGTGGCGCTCACCTCCCTTACATCGAAACAGGGAATGACGCTCGCGG 4550  
QY 3902 AACAAATTCAAAACAGAAAGGCAATCGGTTGCTGCAAAACAGCCACAAAGCAAGCGAGGCTG 3961  
Db 4551 AACAAATTCAAAACAGAAAGGCAATCGGTTGCTGCAAAACAGCCACAAAGCAAGCGAGGCTG 4610  
QY 3962 CTGCTCCGCTGCTGGAAATCCAGTGGGGACCTCGAAGCCTTCTGGGCGAAGCATATGT 4021  
Db 4611 CTGCTCCGCTGCTGGAAATCCAGTGGGGACCTCGAAGCCTTCTGGGCGAAGCATATGT 4670  
QY 4022 GGAATTTTCATCAGCGGATACAAATATTAGCAGGCTTGTCCACTCTGCTGCGCAACCCCG 4081  
Db 4671 GGAATTTTCATCAGCGGATACAAATATTAGCAGGCTTGTCCACTCTGCTGCGCAACCCCG 4730  
QY 4082 CGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCCCGCTCACACCCCAACATA 4141  
Db 4731 CGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCCCGCTCACACCCCAACATA 4790  
QY 4142 CCCTCCTGTTTAAACATCTCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCTG 4201  
Db 4791 CCCTCCTGTTTAAACATCTCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGCGCTG 4850  
QY 4202 CTTCTGCTTTCTGTAGGCGCGCATCGCTGAGCGGCTGTGGCAGCATAGGCTTCGGGA 4261  
Db 4851 CTTCTGCTTTCTGTAGGCGCGCATCGCTGAGCGGCTGTGGCAGCATAGGCTTCGGGA 4910  
QY 4262 AGGTGCTTGTGGATATTTTGGCAGGTTATGGACAGGGGTGGCAGGCGCGCTCGTGGCCT 4321  
Db 4911 AGGTGCTTGTGGATATTTTGGCAGGTTATGGACAGGGGTGGCAGGCGCGCTCGTGGCCT 4970  
QY 4322 TTAAGTTCATGACGGGGAGATGCCCTCCACCGAGACCTGGTTAACTTACTCCTCTCTA 4381  
Db 4971 TTAAGTTCATGACGGGGAGATGCCCTCCACCGAGACCTGGTTAACTTACTCCTCTCTA 5030  
QY 4382 TCCTCTCCCTGCGCCCTAGTCTGCGGGCTGTGTGCGCAGCATACTCGCTCGGCACG 4441  
Db 5031 TCCTCTCCCTGCGCCCTAGTCTGCGGGCTGTGTGCGCAGCATACTCGCTCGGCACG 5090  
QY 4442 TGGGCCCCAGGGGAGGGGCTGTGACGTGGATGAACCGGCTGTAGCGTTTCGCTTCGCGGG 4501  
Db 5091 TGGGCCCCAGGGGAGGGGCTGTGACGTGGATGAACCGGCTGTAGCGTTTCGCTTCGCGGG 5150  
QY 4502 GTAACACAGTCTCCCGACGCACTATGTGCTGAGAGCGACGCTGACGACGCTGTCTACTC 4561  
Db 5151 GTAACACAGTCTCCCGACGCACTATGTGCTGAGAGCGACGCTGACGACGCTGTCTACTC 5210  
QY 4562 AGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGGATCAACG 4621  
Db 5211 AGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGGATCAACG 5270  
QY 4622 AGGACTGCTCCAGCCCATGCTCCGGCTCGTGGCTAAAGAGATGTTTGGGATGGATATGCA 4681  
Db 5271 AGGACTGCTCCAGCCCATGCTCCGGCTCGTGGCTAAAGAGATGTTTGGGATGGATATGCA 5330  
QY 4682 CGGTGTGACTGATTTCAAGACCTGCTCAGTCCAAAGCTCTCGCGCGGATTCGCGGGAG 4741  
Db 5331 CGGTGTGACTGATTTCAAGACCTGCTCAGTCCAAAGCTCTCGCGCGGATTCGCGGGAG 5390  
QY 4742 TCCCTCTTCTCATGTCAACGCTGTACAAGGAGTCTTGGCGGGCGACGCGCATCATGC 4801  
Db 5391 TCCCTCTTCTCATGTCAACGCTGTACAAGGAGTCTTGGCGGGCGACGCGCATCATGC 5450  
QY 4802 AAAACCACTGCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCCATGAGGA 4861  
Db 5451 AAAACCACTGCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCCATGAGGA 5510  
QY 4862 TCGTGGGCGCTAGGACCTGTAGTAAACGTGGCATGGAACATTTCCCATTTAAACGGTACA 4921  
Db 5511 TCGTGGGCGCTAGGACCTGTAGTAAACGTGGCATGGAACATTTCCCATTTAAACGGTACA 5570  
QY 4922 CCACGGGCGCTCAGACGCGCTCCCGGCGCAATTTCTTAGGCGCTGCTGGCGGTGG 4981



```

QY 7142 ACCCCACCCCGCTTGGCGGCTGGTGGGAGACAGTACACACTCCAGTCAATT 7201
Db 7791 ACCCCACCCCGCTTGGCGGCTGGTGGGAGACAGTACACACTCCAGTCAATT 7850
QY 7202 CCTGGCTAGGCAACATCATGATGATGCGGCCACCTTGTGGCAAGGATGATCCTGATGA 7261
Db 7851 CCTGGCTAGGCAACATCATGATGATGCGGCCACCTTGTGGCAAGGATGATCCTGATGA 7910
QY 7262 CTGATTTCTTCTCCATCTCTAGCTAGCTAGGCAACAACTTGAATAAGCCCTAGATGTCAGA 7321
Db 7911 CTGATTTCTTCTCCATCTCTAGCTAGGCAACAACTTGAATAAGCCCTAGATGTCAGA 7970
QY 7322 TCTACGGGGCGCTGTACTCCATTTAGGACCACTTGACCTACCTCAGATCATTCAGGACTCC 7381
Db 7971 TCTACGGGGCGCTGTACTCCATTTAGGACCACTTGACCTACCTCAGATCATTCAGGACTCC 8030
QY 7382 ATGGCCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTT 7441
Db 8031 ACGGCCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCTT 8090
QY 7442 CATGCCCTCAGAAACTTGGGGTACCGCCCTTGCAGTCTGAGACATCGGGCCAGAAAGTG 7501
Db 8091 CATGCCCTCAGAAACTTGGGGTACCGCCCTTGCAGTCTGAGACATCGGGCCAGAAAGTG 8150
QY 7502 TCCGCGCTAGGCTACTCTCCAGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTCA 7561
Db 8151 TCCGCGCTAGGCTACTCTCCAGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTCA 8210
QY 7562 ACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCCCTCCAGTTGGATT 7621
Db 8211 ACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAAATCCCGGCTGCCCTCCAGTTGGATT 8270
QY 7622 TATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAAGCCTGTCTCGTG 7681
Db 8271 TATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAAGCCTGTCTCGTG 8330
QY 7682 CCGACCCCGCTGTTCACTGCTGGTGCCTACTCTCTACTCTCTGTTAGGGGTAGGCACTATC 7741
Db 8331 CCGACCCCGCTGTTCACTGCTGGTGCCTACTCTCTACTCTCTGTTAGGGGTAGGCACTATC 8390
QY 7742 TACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCGCAATAGGCCCATCTCTGTTTTT 7801
Db 8391 TACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCGCAATAGGCCCATCTCTGTTTTT 8450
QY 7802 TTCCC-----TTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCT 7857
Db 8451 TTCCC-----TTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCC 8510
QY 7858 TTTTTTTTCTCTTTTTTTCTTTTTTTCTTTTTTTCTTTTGGTGGCTCCATCTTAGCCCTAGTCAC 7917
Db 8511 TTTTTTTTCTCTTTTTTTCTTTTTTTCTTTTTTTCTTTTTTTGGTGGCTCCATCTTAGCCCTAGTCAC 8570
QY 7918 GGCTAGCTGTGAAGGTCCTGTAGCCGCTTGAAGCTGACGAGAGAGTCTGATATGCGCCTCTC 7977
Db 8571 GGCTAGCTGTGAAGGTCCTGTAGCCGCTTGAAGCTGACGAGAGAGTCTGATGACTGAGAGTCTG 8630
QY 7978 TGCAGATCAAGT 7989
Db 8631 TGCAGATCAAGT 8642

```

```

RESULT 13
US-10-029-907-6
; Sequence 6, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857

```

```

; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6
; LENGTH: 8638
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)... (8407)
US-10-029-907-6

Query Match      90.7%; Score 7244; DB 4; Length 8638;
Best Local Similarity 92.0%; Pred. No. 0;
Matches 7957; Conservative 0; Mismatches 20; Indels 671; Gaps 3;

QY 2 CCAGCCCCCGGATTTGGGGGGCGACACTCCACCATAGATCACTCCCCCTGTGAGGAACTACTGT 61
Db 2 CCAGCCCCCGGATTTGGGGGGCGACACTCCACCATAGATCACTCCCCCTGTGAGGAACTACTGT 61
QY 62 CTTTACGCAAGAAAGCGTCTAGCCATAGCGCTTAGTATGAGTGTGCTGAGCCTCCAGGACC 121
Db 62 CTTTACGCAAGAAAGCGTCTAGCCATAGCGCTTAGTATGAGTGTGCTGAGCCTCCAGGACC 121
QY 122 CCCCCTCCCGGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTCGACGG 181
Db 122 CCCCCTCCCGGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTCGACGG 181
QY 182 ACGACCCGGTCTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGGCGTCCCGCCG 241
Db 182 ACGACCCGGTCTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTTGGGCGTCCCGCCG 241
QY 242 CGAGACTGTAGCCGAGTGTGTGGTTCGGAAGAGCCCTTGTGTAATGCTGCTGATAGGG 301
Db 242 CGAGACTGTAGCCGAGTGTGTGGTTCGGAAGAGCCCTTGTGTAATGCTGCTGATAGGG 301
QY 302 TGTGTCGGAGTCCCGGGAGGCTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAACC 361
Db 302 TGTGTCGGAGTCCCGGGAGGCTCTGTAGACCGTGCACCATGAGCAGCAATCCTAAACC 361
QY 362 TCAAGAAAAACAAAGGGCGCCATGATTGAACAAAGATGGATTGCACGAGGTTCTCC 421
Db 362 TCAAGAAAAACAAAGGGCGCCATGATTGAACAAAGATGGATTGCACGAGGTTCTCC 421
QY 422 GGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAGACAATTCGGCTGCTC 481
Db 422 GGCGCTTGGTGGAGAGGCTATTTCGGCTATGACTGGGCAACAGACAATTCGGCTGCTC 481
QY 482 TGATGCGCGCTGTTCCGGCTGTACGCGAGGGGGCGCCGGTCTTTTCTCAAGACCGA 541
Db 482 TGATGCGCGCTGTTCCGGCTGTACGCGAGGGGGCGCCGGTCTTTTCTCAAGACCGA 541
QY 542 CCTGTCCGGTGCCTTGAATGAACTGCAGGACGAGGCGAGCGCGCTATCGTGGCTGGCCAC 601
Db 542 CCTGTCCGGTGCCTTGAATGAACTGCAGGACGAGGCGAGCGCGCTATCGTGGCTGGCCAC 601
QY 602 GACGGGCTTCTTGGCAGCTGTGCTGAGCTGCTGAGCGGGAAGGAGACTGGCT 661
Db 602 GACGGGCTTCTTGGCAGCTGTGCTGAGCTGCTGAGCGGGAAGGAGACTGGCT 661
QY 662 GCTATTGGGGGAAGTCCCGGGGAGGATCTCTGTGATCTCACTTGTCTCTCCCGAGAA 721
Db 662 GCTATTGGGGGAAGTCCCGGGGAGGATCTCTGTGATCTCACTTGTCTCTCCCGAGAA 721
QY 722 AGTATCCATCATGGCTGATGCAATTCGGCGGCTGCAATACGCTTGTATCCGCTACCTGCC 781
Db 722 AGTATCCATCATGGCTGATGCAATTCGGCGGCTGCAATACGCTTGTATCCGCTACCTGCC 781
QY 782 ATTCGACCAACAGCGAAACATCGATTCGAGCGAGCAAGTACTCGGATGAAAGCGCGTCT 841
Db 782 ATTCGACCAACAGCGAAACATCGATTCGAGCGAGCAAGTACTCGGATGAAAGCGCGTCT 841
QY 842 TGTGATCAGATGATCTGACGAGCAATCAGGGGCTCGCGCCAGCCGAACTGTTCCG 901

```

Db 842 TGTGATCAGGATGATCTGACGAGAGCATCAGGGGCTCGCGCAGCCGAACTGTTCCG 901  
Qy 902 CAGGCTCAAGGGCGGATGCCGACGCGAGGATCTCGTGTGACCCATGGGATGCCCTG 961  
Db 902 CAGGCTCAAGGGCGGATGCCGACGCGAGGATCTCGTGTGACCCATGGGATGCCCTG 961  
Qy 962 CTTGGCGAATATCATGGTGGAAATGGCCGCTTTCTGGATTCATCATGCTGGCGGCT 1021  
Db 962 CTTGGCGAATATCATGGTGGAAATGGCCGCTTTCTGGATTCATCATGCTGGCGGCT 1021  
Qy 1022 GGGTGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCCGTTGATATTCTGAAGAGCT 1081  
Db 1022 GGGTGTGGCGGACCGCTATCAGACATAGCGTTGGCTACCCGTTGATATTCTGAAGAGCT 1081  
Qy 1082 TGGCGGCGAATGGGTGACCGCTTCTCGTGTGTTACGGTATCGCGCTGCCGATTCGCA 1141  
Db 1082 TGGCGGCGAATGGGTGACCGCTTCTCGTGTGTTACGGTATCGCGCTGCCGATTCGCA 1141  
Qy 1142 GGGCATGCGCTTCTATGCGCTTCTTGACGAGTTCTTCTGAGTT-----TAAAC 1189  
Db 1142 GGGCATGCGCTTCTATGCGCTTCTTGACGAGTTCTTCTGAGTT-----TAAAC 1201  
Qy 1190 AGACCAACAGGTTTCCCTCTAGCGGATCAATTCGCCCTTCTCCCTCCGCCCCCTTA 1249  
Db 1202 AGACCAACAGGTTTCCCTCTAGCGGATCAATTCG-----CCCCCCCCCTA 1250  
Qy 1250 AGTTACTGGCGGACCGCTTGGAAATAGGCGGTGTCGTTTGTCTATGTTATTTT 1309  
Db 1251 AGTTACTGGCGGACCGCTTGGAAATAGGCGGTGTCGTTTGTCTATGTTATTTT 1310  
Qy 1310 CCACCATATGCGCTTCTTGGCAATGTGAGGCGCGGAAACCTGCGCTGCTTCTTGA 1369  
Db 1311 CCACCATATGCGCTTCTTGGCAATGTGAGGCGCGGAAACCTGCGCTGCTTCTTGA 1370  
Qy 1370 CGAGCATTCCTAGGGTCTTTCCCTCTGCCAAAGGAATGAAGTCTGTTGAATGCG 1429  
Db 1371 CGAGCATTCCTAGGGTCTTTCCCTCTGCCAAAGGAATGAAGTCTGTTGAATGCG 1430  
Qy 1430 TGAAGGAGCAGTTCTCTGGAAGTTCTTGAAGCAAAACACGCTGTGAGGACCTTT 1489  
Db 1431 TGAAGGAGCAGTTCTCTGGAAGTTCTTGAAGCAAAACACGCTGTGAGGACCTTT 1490  
Qy 1490 GCAGGACGCGAACCCTGCGACAGTGTCTGCGGCCAAAGCCACGCTGAT 1549  
Db 1491 GCAGGACGCGAACCCTGCGACAGTGTCTGCGGCCAAAGCCACGCTGAT 1550  
Qy 1550 AAGATACACCTGCAAGGCGGCACAAACCCAGTGCACGTTGTGATGTTG 1609  
Db 1551 AAGATACACCTGCAAGGCGGCACAAACCCAGTGCACGTTGTGATGTTG 1610  
Qy 1610 AAGAGTCAATGGCTCTCTCAAGGTATTCAAGGGGCTGAAGATGCCGAGAG 1669  
Db 1611 AAGAGTCAATGGCTCTCTCAAGGTATTCAAGGGGCTGAAGATGCCGAGAG 1670  
Qy 1670 TACCCATTGTATGGATCTGATCTGGGCTCTGCTGACATGCTTTACATGTTTGT 1729  
Db 1671 TACCCATTGTATGGATCTGATCTGGGCTCTGCTGACATGCTTTACATGTTTGT 1730  
Qy 1730 CGAGTTTAAACCTCTAGGCCCCCGAAACGCGGGACGTTTCTTTTGAAGAAC 1789  
Db 1731 CGAGTTTAAACCTCTAGGCCCCCGAAACGCGGGACGTTTCTTTTGAAGAAC 1790  
Qy 1790 ACGATAATACCATG-----1803  
Db 1791 ACGATAATACCATGACCGGAGATGGCAGATCTGCGGAGCGGCTTTTCGTAGGTC 1850  
Qy 1804 -----1803  
Db 1851 TGATACTCTTGACCTTGTACCGCACAATAAGCTGTTCTCGTAGGCTCATATGGTGT 1910  
Qy 1804 -----1803

Db 1911 TACAATATTTTATCACCAGGCGGAGGCACACTTGTCAAGTGTGGATCCCCCCTCAACG 1970  
Qy 1804 -----1803  
Db 1971 TTGCGGGGGCGCGATGCCGTATCTCTCACGTGCGGATCCACCAGAGCTAATCT 2030  
Qy 1804 -----1803  
Db 2031 TTACCATCAACAAATCTTGTCTGCCATACTCGTCCACTCATGTGTCTCCAGGCTGGTA 2090  
Qy 1804 -----1803  
Db 2091 TAAACAAAGTGGCGTACTTTCGTGCGGCACACGGGCTCATTCGTGATCATCTGTGTC 2150  
Qy 1804 -----1803  
Db 2151 GGAAGGTTGCTGGGGTCAATATGTCCAAATGGCTCTCATGAATGGCTGGCGCACTGACAG 2210  
Qy 1804 -----1803  
Db 2211 GTACGTACGTTTATGACCATCTCACCCACTCGGGACTGGGCCACGCGGCTTACGAG 2270  
Qy 1804 -----1803  
Db 2271 ACCTTGGGTGCGAGTTGAGCCCGTCTTCTCTGATATGGAGCAAGGTTATCACCT 2330  
Qy 1804 -----1803  
Db 2331 GGGGGGCGAGACACCGCGGTGTGGGACATCATCTTGGGCTGCGCGTCTCGGCCGCA 2390  
Qy 1804 -----1803  
Db 2391 GGGGGAGGAGATACATCTGGGACCGGCAGACAGCTTGAAGGGCAGGGGTGGGACTCC 2450  
Qy 1804 --GCCCTATTACGCCCTACTCCCAACAGACGCGAGGCTTCTTGGCTGCATCATCACTA 1861  
Db 2451 TCGCGCTTATTACGCCCTACTCCCAACAGACGCGAGGCTTCTTGGCTGCATCATCACTA 2510  
Qy 1862 GCTCACAGGCGGACAGGACACAGTTCGAGGGGAGGTCCAAAGTGTCTCCACGGCAA 1921  
Db 2511 GCTCACAGGCGGACAGGAAACAGGTTCGAGGGGAGGTCCAAAGTGTCTCCACGGCAA 2570  
Qy 1922 CACAATCTTCTGCGGACTGCGTCAATGGGTGTGTTGGACTGTCTATCATGTGCGG 1981  
Db 2571 CACAATCTTCTGCGGACTGCGTCAATGGGTGTGTTGGACTGTCTATCATGTGCGG 2630  
Qy 1982 GCTCAAGACCTTTCGCGGCCAAAGGGCCAAATCACCACAAATGTACACCAATGTGGACC 2041  
Db 2631 GCTCAAGACCTTTCGCGGCCAAAGGGCCAAATCACCACAAATGTACACCAATGTGGACC 2690  
Qy 2042 AGGACTCTGTCGGCTGGCAGGCGCCCGGGGCGGTTCTTGTGACCATGCACTTGG 2101  
Db 2691 AGGACTCTGTCGGCTGGCAGGCGCCCGGGGCGGTTCTTGTGACCATGCACTTGG 2750  
Qy 2102 CGAGCTCGGACCTTTACTTGTGTCAGGCGATGCCGATGTCAATCCGGTGCSCGCGGG 2161  
Db 2751 CGAGCTCGGACCTTTACTTGTGTCAGGCGATGCCGATGTCAATCCGGTGCSCGCGGG 2810  
Qy 2162 GCGACAGAGGGGAGCTTCTCTCCCGAGCGCGTCTCTTCTTGAAGGCTCTTGG 2221  
Db 2811 GCGAGGAGGGGAGGCTTCTCTCCCGAGCGCGTCTCTTCTTGAAGGCTCTTGG 2870  
Qy 2222 GGGTCTCACTCTGCGCCCTCGGGGACGCTGTGGGATCTTTCGGGCTGCGGTGCA 2281  
Db 2871 GGGTCTCACTCTGCGCCCTCGGGGACGCTGTGGGATCTTTCGGGCTGCGGTGCA 2930  
Qy 2282 CCCAGGAGGTTGCGAAGGCGGTGGAATTTGTACCCGTGAGTCTATGAAACACCATATGC 2341  
Db 2931 CCCAGGAGGTTGCGAAGGCGGTGGAATTTGTACCCGTGAGTCTATGAAACACCATATGC 2990  
Qy 2342 GGTCCCCGGTCTTACCGACAACTCTGTCCTTCCGGCGTACCGGACATTCAGGTTGG 2401  
Db 2991 GGTCCCCGGTCTTACCGACAACTCTGTCCTTCCGGCGTACCGGACATTCAGGTTGG 3050



2402 QY CCCATCTACAGCGCCCTACTGTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCGAG 2461  
Db |||||  
3051 QY CCCATCTACAGCGCCCTACTGTGTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCGAG 3110  
Db |||||  
2462 QY CCCAAGGGTATAAGGTGCTTGTCTGAACCCGTCGCTCGCGGCACCCCTAGGTTCGCGG 2521  
Db |||||  
3111 QY CCCAAGGGTATAAGGTGCTTGTCTGAACCCGTCGCTCGCGGCACCCCTAGGTTCGCGG 3170  
Db |||||  
2522 QY CGTATATGTCTAAGGCACATGTGTATCGACCTTAACATCAGAACCAGGGTAAAGACCATCA 2581  
Db |||||  
3171 QY CGTATATGTCTAAGGCACATGTGTATCGACCTTAACATCAGAACCAGGGTAAAGACCATCA 3230  
Db |||||  
2582 QY CACGGGTGCCCCCATCAAGTACTCCACCTATCGGCAAGTTCTTGGCAGCGTGGTGTCT 2641  
Db |||||  
3231 QY CACGGGTGCCCCCATCAAGTACTCCACCTATCGGCAAGTTCTTGGCAGCGTGGTGTCT 3290  
Db |||||  
2642 QY CTGGGGCGGCTATGACATCATATAATATGTATGATGAGTGCACCTCAACTGACTCGACCACTA 2701  
Db |||||  
3291 QY CTGGGGCGGCTATGACATCATATAATATGTATGATGAGTGCACCTCAACTGACTCGACCACTA 3350  
Db |||||  
2702 QY TCCTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGCGGCTGGAGCGGCACTCGTCTGTC 2761  
Db |||||  
3351 QY TCCTGGGCATCGGCACAGTCTCTGGACCAAGCGGAGCGGCTGGAGCGGCACTCGTCTGTC 3410  
Db |||||  
2762 QY TCGCCACCGCTACGCTCCGGGATCGGTACCGTGCACATCCAAACATCGAGGAGTGG 2821  
Db |||||  
3411 QY TCGCCACCGCTACGCTCCGGGATCGGTACCGTGCACATCCAAACATCGAGGAGTGG 3470  
Db |||||  
2822 QY CTCTGTCCAGCTGGAGAAATCCCTTTTATGCAAGAGCCATCCCATCGAGACCATCA 2881  
Db |||||  
3471 QY CTCTGTCCAGCTGGAGAAATCCCTTTTATGCAAGAGCCATCCCATCGAGACCATCA 3530  
Db |||||  
2882 QY AGGGGGGGAGGCACTCATTTTCTGCCATTCGAAGAGAAATGTGATGAGCTCGCGCGA 2941  
Db |||||  
3531 QY AGGGGGGGAGGCACTCATTTTCTGCCATTCGAAGAGAAATGTGATGAGCTCGCGCGA 3590  
Db |||||  
2942 QY AGCTGTCCGGCTCGGACTCAATGTCTGATAGCATATTAACGGGGCTTGATGATCGTCA 3001  
Db |||||  
3591 QY AGCTGTCCGGCTCGGACTCAATGTCTGATAGCATATTAACGGGGCTTGATGATCGTCA 3650  
Db |||||  
3002 QY TACCAACTAGCGAGAGCTCATTTGCTAGCAAGGAGCGCTTAAAGACGGGCTTTACCG 3061  
Db |||||  
3651 QY TACCAACTAGCGAGAGCTCATTTGCTAGCAAGGAGCGCTTAAAGACGGGCTTTACCG 3710  
Db |||||  
3062 QY GCGATTTGCACTCAGTGATCGACTGCAATATCATGTGTCAACCAGACAGTCACTTCAGCC 3121  
Db |||||  
3711 QY GCGATTTGCACTCAGTGATCGACTGCAATATCATGTGTCAACCAGACAGTCACTTCAGCC 3770  
Db |||||  
3122 QY TGGACCGGACCTTCAACATTTAGACAGCAGCAGCTGCGCAAGAGCGGCTGACGCTCGC 3181  
Db |||||  
3771 QY TGGACCGGACCTTCAACATTTAGACAGCAGCAGCTGCGCAAGAGCGGCTGACGCTCGC 3830  
Db |||||  
3182 QY AGCGGGGAGGAGGAGTGTAGGGGAGGATGGGCAATTTACAGTTTGTGACTCCAGGAG 3241  
Db |||||  
3831 QY AGCGGGGAGGAGGAGTGTAGGGGAGGATGGGCAATTTACAGTTTGTGACTCCAGGAG 3890  
Db |||||  
3242 QY AACGGGCTCGGGCATTTGCAATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGTG 3301  
Db |||||  
3891 QY AACGGGCTCGGGCATTTGCAATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGTG 3950  
Db |||||  
3302 QY CTTGGTACGAGCTACGCGCGCGAGACCTCAGTTAGTGTGCGGCTTACCTAAACACAC 3361  
Db |||||  
3951 QY CTTGGTACGAGCTACGCGCGCGAGACCTCAGTTAGTGTGCGGCTTACCTAAACACAC 4010  
Db |||||  
3362 QY CAGGGTTGCCGCTGCGCAGGACCATCTGAGTTCTGGAGAGCGTCTTTACAGGCGCTCA 3421  
Db |||||  
4011 QY CAGGGTTGCCGCTGCGCAGGACCATCTGAGTTCTGGAGAGCGTCTTTACAGGCGCTCA 4070  
Db |||||  
3422 QY CCCACATAGAGCGCCCATTTCTGTGCCAGACTAAGCAGGAGGAGACAACTTCCCTTACC 3481  
Db |||||  
4071 QY CCCACATAGAGCGCCCATTTCTGTGCCAGACTAAGCAGGAGGAGACAACTTCCCTTACC 4130  
Db |||||

3482 QY TGGTAGCATACAGGCTACGGTGTGCGCAGGCTCAGGCTCCACCTCCATCGTGGAGC 3541  
Db |||||  
4131 QY TGGTAGCATACAGGCTACGGTGTGCGCAGGCTCAGGCTCCACCTCCATCGTGGAGC 4190  
Db |||||  
3542 QY AAATGTGGAAGTGTCTCATACAGCTAAAGCTACGCTGACCGGCCAAACGCCCTGTGT 3601  
Db |||||  
4191 QY AAATGTGGAAGTGTCTCATACAGCTAAAGCTACGCTGACCGGCCAAACGCCCTGTGT 4250  
Db |||||  
3602 QY ATAGGCTGGAGCGGTTCAAAAGAGGTTACTACCAACACCCCATAAACAAATACATCA 3661  
Db |||||  
4251 QY ATAGGCTGGAGCGGTTCAAAAGAGGTTACTACCAACACCCCATAAACAAATACATCA 4310  
Db |||||  
3662 QY TGGCATGATGTCGGCTGACCTGGAGTGTGTCAGGAGCCTGGGTGCTGGTAGGCGGAG 3721  
Db |||||  
4311 QY TGGCATGATGTCGGCTGACCTGGAGTGTGTCAGGAGCCTGGGTGCTGGTAGGCGGAG 4370  
Db |||||  
3722 QY TCCTAGCAGCTCTGGCGGGGTATTGCTGACAAACAGCAGCGTGGTCAATTTGTTGGGAGGA 3781  
Db |||||  
4371 QY TCCTAGCAGCTCTGGCGGGTATTGCTGACAAACAGCAGCGTGGTCAATTTGTTGGGAGGA 4430  
Db |||||  
3782 QY TCACTTTGTCGGAAGAGCGGCCCATCATTCGCCAGCAGGGAAGTCTTTTACCGGGAGTTCG 3841  
Db |||||  
4431 QY TCACTTTGTCGGAAGAGCGGCCCATCATTCGCCAGCAGGGAAGTCTTTTACCGGGAGTTCG 4490  
Db |||||  
3842 QY ATGAGATGGAAGAGTGGCTCACCTCCCTTACATCGAAACAGGAATGCAGCTCGCCG 3901  
Db |||||  
4491 QY ATGAGATGGAAGAGTGGCTCACCTCCCTTACATCGAAACAGGAATGCAGCTCGCCG 4550  
Db |||||  
3902 QY AACAAATCAAAACAGAGGCAATCGGGTTGCTGCAAAACAGCACCAAGCAAGCGGAGCTG 3961  
Db |||||  
4551 QY AACAAATCAAAACAGAGGCAATCGGGTTGCTGCAAAACAGCACCAAGCAAGCGGAGCTG 4610  
Db |||||  
3962 QY CTGCTCCGCTGTGGAAATCCAGTGGCGGACCTCGAAAGCTTCTGGGCGAAGCATATGT 4021  
Db |||||  
4611 QY CTGCTCCGCTGTGGAAATCCAGTGGCGGACCTCGAAAGCTTCTGGGCGAAGCATATGT 4670  
Db |||||  
4022 QY GGAATTTATCAGCGGATACAAATATTAGCAGGCTGTCCACTCTGCTGGCAACCCCG 4081  
Db |||||  
4671 QY GGAATTTATCAGCGGATACAAATATTAGCAGGCTGTCCACTCTGCTGGCAACCCCG 4730  
Db |||||  
4082 QY CGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCCGCTCACCAACCAACATA 4141  
Db |||||  
4731 QY CGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCCGCTCACCAACCAACATA 4790  
Db |||||  
4142 QY CCCTCTGTTTAACTATCTGCGGGGATGCGTGGCGGCCAACTTGTCTCTCCAGGCGCTG 4201  
Db |||||  
4791 QY CCCTCTGTTTAACTATCTGCGGGGATGCGTGGCGGCCAACTTGTCTCTCCAGGCGCTG 4850  
Db |||||  
4202 QY CTTCTGCTTTCTAGGCGCGGCATCGCTGAGCGGCTGTTGGCAGCATAGGCTTCGGGA 4261  
Db |||||  
4851 QY CTTCTGCTTTCTAGGCGCGGCATCGCTGAGCGGCTGTTGGCAGCATAGGCTTCGGGA 4910  
Db |||||  
4262 QY AGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGGCTCGTGGCCT 4321  
Db |||||  
4911 QY AGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGGTGGCAGGCGGCTCGTGGCCT 4970  
Db |||||  
4322 QY TTAAGTCTAGAGCGGCGAGATCCCTTCCACCGAGACCTGGTTAACTACTCCCTGCTA 4381  
Db |||||  
4971 QY TTAAGTCTAGAGCGGCGAGATCCCTTCCACCGAGACCTGGTTAACTACTCCCTGCTA 5030  
Db |||||  
4382 QY TCCTCTCCCTGGCGGCTAGTCTGGGTGCTGCGCAGCATACTCCGTCGGCAGC 4441  
Db |||||  
5031 QY TCCTCTCCCTGGCGGCTAGTCTGGGTGCTGCGCAGCATACTCCGTCGGCAGC 5090  
Db |||||  
4442 QY TGGGCGCCAGGGAGGGGCTGTGCAATGGAATGAACCGGCTGATAGGCTTCGCTTCGCGG 4501  
Db |||||  
5091 QY TGGGCGCCAGGGAGGGGCTGTGCAATGGAATGAACCGGCTGATAGGCTTCGCTTCGCGG 5150  
Db |||||  
4502 QY GTAAACGAGCTCTCCCGACGCACTATGTGCTGAGAGCGACGCTGCAGACGCTGTCACTC 4561  
Db |||||  
5151 QY GTAAACGAGCTCTCCCGACGCACTATGTGCTGAGAGCGACGCTGCAGACGCTGTCACTC 5210  
Db |||||  
4562 QY AGATCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCCACGATGAGCAACG 4621  
Db |||||

Db	5211	AGATCCCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCAACAGTGGATCAACG	5270
Qy	4622	AGGACTGCTCCACGCCATGCTCCGGCTCGTGCCATAAGAGATGTTTGGATTTGATATGCA	4681
Db	5271	AGGACTGCTCCACGCCATGCTCCGGCTCGTGCCATAAGAGATGTTTGGATTTGATATGCA	5330
Qy	4682	CGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCTGCCCGGATTTGCCGGAG	4741
Db	5331	CGGTGTTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGCTCCTGCCCGGATTTGCCGGAG	5390
Qy	4742	TCCCTTCTTCTCATGTCAACGTGGGTACAAAGGAGTCTTGGCGGGGCGACGGGCATCATGC	4801
Db	5391	TCCCTTCTTCTCATGTCAACGTGGGTACAAAGGAGTCTTGGCGGGGCGACGGGCATCATGC	5450
Qy	4802	AAACCACTGCCATGTGGACACAGATCAACGGACATGTGMAAAAGGTTTCCATGAGGA	4861
Db	5451	AAACCACTGCCATGTGGACACAGATCAACGGACATGTGMAAAAGGTTTCCATGAGGA	5510
Qy	4862	TCGTGGGGCCCTAGGACCTGTAGTAAACGTGGCATGGAACATTTCCCAATTAACGCGTACA	4921
Db	5511	TCGTGGGGCCCTAGGACCTGTAGTAAACGTGGCATGGAACATTTCCCAATTAACGCGTACA	5570
Qy	4922	CCACGGGGCCCTGCACGGCCCTCCCGGGGCGCAAAATTAATTAGGGCGCTGTGCGGGGTGG	4981
Db	5571	CCACGGGGCCCTGCACGGCCCTCCCGGGGCGCAAAATTAATTAGGGCGCTGTGCGGGGTGG	5630
Qy	4982	CTGCTGAGGAGTAGCTGGAGGTTACCGGGGTGGGGATTTCCACTACGTACACGGGCATGA	5041
Db	5631	CTGCTGAGGAGTAGCTGGAGGTTACCGGGGTGGGGATTTCCACTACGTACACGGGCATGA	5690
Qy	5042	CCACTGACAAAGTAAAGTGCCCGGTTCAGAGTTCGGGCCCGCAATTTCTTCACAGAAAGTGG	5101
Db	5691	CCACTGACAAAGTAAAGTGCCCGGTTCAGAGTTCGGGCCCGCAATTTCTTCACAGAAAGTGG	5750
Qy	5102	ATGGGGTGGCGTTGCAACAGTACGCTCCACGGTGCAAACCCCTCTCAACGGGAGGAGTCA	5161
Db	5751	ATGGGGTGGCGTTGCAACAGTACGCTCCACGGTGCAAACCCCTCTCAACGGGAGGAGTCA	5810
Qy	5162	CAATTCCTGGTCGGGCTCAATCAATACCTGGTTGGGTACAGCTCCCATCGGACCGCGAAC	5221
Db	5811	CAATTCCTGGTCGGGCTCAATCAATACCTGGTTGGGTACAGCTCCCATCGGACCGCGAAC	5870
Qy	5222	CGGACGTAGCAGTGTCACTTCCATGCTCACCGACCCCTCCCACTTACGGCGGAGACGG	5281
Db	5871	CGGACGTAGCAGTGTCACTTCCATGCTCACCGACCCCTCCCACTTACGGCGGAGACGG	5930
Qy	5282	CTAAGCGTAGGCTGGCCAGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAGC	5341
Db	5931	CTAAGCGTAGGCTGGCCAGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAGC	5990
Qy	5342	TGCTCGGCGCTTCTTGAAGGCAATGCACTACCGGTATGACTCCCGGACGCTGACC	5401
Db	5991	TGCTCGGCGCTTCTTGAAGGCAATGCACTACCGGTATGACTCCCGGACGCTGACC	6050
Qy	5402	TCATCGAGGCCAACCTCTCTGTGGCGCAGAGATGGCGGGAAACATCACCCGGTGGAGT	5461
Db	6051	TCATCGAGGCCAACCTCTCTGTGGCGCAGAGATGGCGGGAAACATCACCCGGTGGAGT	6110
Qy	5462	CAGAAATAAGGTAGTAATTTGGACTTTTCGAGCCGCTCCAAGCGGAGGAGATGAGA	5521
Db	6111	CAGAAATAAGGTAGTAATTTGGACTTTTCGAGCCGCTCCAAGCGGAGGAGATGAGA	6170
Qy	5522	GGGAGTAGTCGTTCCGGCGGAGATCCTGGGAGGTCCAGGAAATTTCCCTCGAGCGATGC	5581
Db	6171	GGGAGTAGTCGTTCCGGCGGAGATCCTGGGAGGTCCAGGAAATTTCCCTCGAGCGATGC	6230
Qy	5582	CCATATGGGCACGCCCGGATTACAAACCTCCACTGTTAGAGTCTCTGGAAGGACCCGGACT	5641
Db	6231	CCATATGGGCACGCCCGGATTACAAACCTCCACTGTTAGAGTCTCTGGAAGGACCCGGACT	6290
Qy	5642	ACGTCCCTCCAGTGGTACACGGGTGTCATTTCCGCGCTGCAGAGGCCCTTCGATACCAC	5701

[illegible]

RESULT 14

```

US-10-029-907-5
; Sequence 5, Application US/10029907
; Patent No. 6706874
; GENERAL INFORMATION:
; APPLICANT: BOHRINGER INGELHEIM (CANADA) LTD.
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM
; FILE OF INVENTION: HEPATITIS C VIRUS
; FILE REFERENCE: 13/083
; CURRENT APPLICATION NUMBER: US/10/029,907
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: 60/257,857
; PRIOR FILING DATE: 2000-12-22
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 8648
; TYPE: DNA
; ORGANISM: HCV
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1802)...(8407)
US-10-029-907-5

```

Query Match 90.5%; Score 7228.2; DB 4; Length 8648;  
Best Local Similarity 91.9%; Pred. No. 0;  
Matches 7960; Conservative 0; Mismatches 18; Indels 691;

Qy	1	GCACGCCCCGATTGGGGCGACACTCCACCATGATCACTCCCTCTGTAGGAACTACTCG	60
Db	1	GCACGCCCCGATTGGGGCGACACTCCACCATGATCACTCCCTCTGTAGGAACTACTCG	60
Qy	61	TCCTTCACGCAGAAACGCTCTAGCCATGCGCTTAGTATGAGTGTCTGTGACGCTCCAGAGC	120
Db	61	TCCTTCACGCAGAAACGCTCTAGCCATGCGCTTAGTATGAGTGTCTGTGACGCTCCAGGAC	120
Qy	121	CCCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGTGAGTACACCGGAATTTGCCAG	180
Db	121	CCCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGTGAGTACACCGGAATTTGCCAG	180
Qy	181	GACGACCGGGTCCTTTCTTGATCAACCCCGCTCAATGCTCGAGATTTGGCGTGCCTCC	240
Db	181	GACGACCGGGTCCTTTCTTGATCAACCCCGCTCAATGCTCGAGATTTGGCGTGCCTCC	240
Qy	241	GCAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAAGCCCTTGTGGTATCTGCTGATAGG	300
Db	241	GCAGACTGCTAGCCGAGTAGTGTGGTTCGCGAAAGCCCTTGTGGTATCTGCTGATAGG	300
Qy	301	GTGCTTTCGGAGTGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC	360
Db	301	GTGCTTTCGGAGTGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC	360
Qy	361	CTCAAGAAAAAACCAGAGGGCGGCCCATGATTGAAACAAGATGGATTGCACGAGGTTCTC	420
Db	361	CTCAAGAAAAAACCAGAGGGCGGCCCATGATTGAAACAAGATGGATTGCACGAGGTTCTC	420
Qy	421	CGGCCGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAAGCAATTCGGCTGCT	480
Db	421	CGGCCGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAAGCAATTCGGCTGCT	480

QY 481 CTGATCGCGCGTGTCCGGCTGTCTAGCGAGGGGCGCCGGTCTCTTTTGTCAAGACCG 540  
 Db |||||  
 QY 481 CTGATCGCGCGTGTCCGGCTGTCTAGCGAGGGGCGCCGGTCTCTTTTGTCAAGACCG 540  
 Db |||||  
 QY 541 ACTGTCCGGTGCCTTGAATGAATCTGACGAGAGGCGAGCGGCTATCGTGGCTGGCCA 600  
 Db |||||  
 QY 541 ACTGTCCGGTGCCTTGAATGAATCTGACGAGAGGCGAGCGGCTATCGTGGCTGGCCA 600  
 Db |||||  
 QY 601 CGACGGCGCTTCTCTGCGCAGCTGTCTGAGCTGTCTCACTGAGCGGGAAGGAGCTGGC 660  
 Db |||||  
 QY 601 CGACGGCGCTTCTCTGCGCAGCTGTCTGAGCTGTCTCACTGAGCGGGAAGGAGCTGGC 660  
 Db |||||  
 QY 661 TGTCTATTGGCGGAAGTGCAGGCGAGGATCTCTCTGTCACTCTCACTCTGCTGCGGAGA 720  
 Db |||||  
 QY 661 TGTCTATTGGCGGAAGTGCAGGCGAGGATCTCTCTGTCACTCTCACTCTGCTGCGGAGA 720  
 Db |||||  
 QY 721 AAGTATCCATCTATGGCTGATGCAATCGCGGCGTGCATACGCTTGTATCGCGCTACCTGCC 780  
 Db |||||  
 QY 721 AAGTATCCATCTATGGCTGATGCAATCGCGGCGTGCATACGCTTGTATCGCGCTACCTGCC 780  
 Db |||||  
 QY 781 CATTCGACCAACGAGGAAACATCGCATCGAGCGACGCTACTCGGATGGAAGCCGGTC 840  
 Db |||||  
 QY 781 CATTCGACCAACGAGGAAACATCGCATCGAGCGACGCTACTCGGATGGAAGCCGGTC 840  
 Db |||||  
 QY 841 TTGTCTGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGCGCCAGCGCACTTTCCG 900  
 Db |||||  
 QY 841 TTGTCTGATCAGGATGATCTGGAAGAGAGCATCAGGGGCTCGCGCCAGCGCACTTTCCG 900  
 Db |||||  
 QY 901 CAGGCTCAAGGCGCGCATGCCGAGCGGAGGATCTCGTGTGACCCATGCGGATGCGCT 960  
 Db |||||  
 QY 901 CAGGCTCAAGGCGCGCATGCCGAGCGGAGGATCTCGTGTGACCCATGCGGATGCGCT 960  
 Db |||||  
 QY 961 GCTTCCCGAATATCATGTGTGGAATAAGCGCGCTTTCTGATTCATCGACTGTGGCCGCG 1020  
 Db |||||  
 QY 961 GCTTCCCGAATATCATGTGTGGAATAAGCGCGCTTTCTGATTCATCGACTGTGGCCGCG 1020  
 Db |||||  
 QY 1021 TGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTTGCTGAAGAGC 1080  
 Db |||||  
 QY 1021 TGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATTTGCTGAAGAGC 1080  
 Db |||||  
 QY 1081 TTGGGGCGAAATGGGGCTGACCGCTTCTCTGTGTTTACGGTATCGCGCTCCGATTCGC 1140  
 Db |||||  
 QY 1081 TTGGGGCGAAATGGGGCTGACCGCTTCTCTGTGTTTACGGTATCGCGCTCCGATTCGC 1140  
 Db |||||  
 QY 1141 AGCGCATCGCTTCTATCGCTTCTTTGACGAGTCTTCTGAGTT-----TAAA 1188  
 Db |||||  
 QY 1141 AGCGCATCGCTTCTATCGCTTCTTTGACGAGTCTTCTGAGTTCTGAGTTCGGGCCAGATGTTAA 1200  
 Db |||||  
 QY 1189 CAGACCAACCGTTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCT 1248  
 Db |||||  
 QY 1201 CAGACCAACCGTTTCCCTCTAGCGGATCAATTCG-----CCCCCCCCCT 1249  
 Db |||||  
 QY 1249 AAGCTTACTTGGCGAAGCGCTTGGAAATAAGCGCGTGTGGCTTGTCTATATGTTATTT 1308  
 Db |||||  
 QY 1250 AAGCTTACTTGGCGAAGCGCTTGGAAATAAGCGCGTGTGGCTTGTCTATATGTTATTT 1309  
 Db |||||  
 QY 1309 TCCACCATATGCGCTTTTGGCAATGTAGGCGCCGAAACCTGCGCTGCTCTTTG 1368  
 Db |||||  
 QY 1310 TCCACCATATGCGCTTTTGGCAATGTAGGCGCCGAAACCTGCGCTGCTCTTTG 1369  
 Db |||||  
 QY 1369 ACGAGCATCTTAGGGGTCTTTCCCTCTCGCAAGGAATGCAAGGTCTGTTGAATGTC 1428  
 Db |||||  
 QY 1370 ACGAGCATCTTAGGGGTCTTTCCCTCTCGCAAGGAATGCAAGGTCTGTTGAATGTC 1429  
 Db |||||  
 QY 1429 GTGAGGAAGAGTCTCTCTGGAAGCTTCTTGAAGACAAACAGTCTGTAGGACCTTT 1488  
 Db |||||  
 QY 1430 GTGAGGAAGAGTCTCTCTGGAAGCTTCTTGAAGACAAACAGTCTGTAGGACCTTT 1489  
 Db |||||  
 QY 1489 TGCAGCAGCGGAACCCCGACCTGCGGACAGTGTCTTGGGCCCAAAAGCCACCTGTA 1548  
 Db |||||  
 QY 1490 TGCAGCAGCGGAACCCCGACCTGCGGACAGTGTCTTGGGCCCAAAAGCCACCTGTA 1549  
 Db |||||  
 QY 1549 TAAGATACCTTGCAAAAGCGGCAACACCCAGTGCCACGTTGTGATGGATGTTGTG 1608  
 Db |||||

Db 1550 TAAGATACACCTTGCAAAAGCGCGCAACCCAGTGCCAGTGTGTGATGGATGTTGTG 1609  
 QY |||||  
 QY 1609 GAAAGAGTCAATGGCTCTCTCAAGCGTATTTCAACAAGGGCTGAAGATGCCAGAAG 1668  
 Db |||||  
 QY 1610 GAAAGAGTCAATGGCTCTCTCAAGCGTATTTCAACAAGGGCTGAAGATGCCAGAAG 1669  
 Db |||||  
 QY 1669 GTACCCCATTTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTGTAG 1728  
 Db |||||  
 QY 1670 GTACCCCATTTATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTGTAG 1729  
 Db |||||  
 QY 1729 TCGAGTTAAAAAAGCTCTAGGCCCGCCGAAACAAGGAGCGTGGTTTCTTTTGA AAA 1788  
 Db |||||  
 QY 1730 TCGAGTTAAAAAAGCTCTAGGCCCGCCGAAACAAGGAGCGTGGTTTCTTTTGA AAA 1789  
 Db |||||  
 QY 1789 CAGGATATACCATG----- 1803  
 Db |||||  
 QY 1790 CAGGATATACCATGAGACGGGAGATGGCAGCATCGTGGAGGCGCGTTCGTAGGT 1849  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 1850 CTGATACTCTTGACCTTGTACCGGACATATAAGCTGTTCTCGCTAGGCTCATATGTTG 1909  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 1910 TTACAATAATTTTATCACAGGCGGAGGCACACTTGCAAGTGTGGATCCCCCCTCAAC 1969  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 1970 GTTCGGGGGGCGCGAGTGCCTCATCTCTCAAGTGTGGCGGATCCACAGAGCTAATC 2029  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 2030 TTTACCATCACAAATCTTGTCTCGCCATCTCGGTCCACTCATGTGTCTCCAGGCTGTT 2089  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 2090 ATACCAAGTGCCTGCTTCTGCGCGCACAGGCTCATCTGTCATGCTCATGCTGTTG 2149  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 2150 CGAAGGTTGCTGGGGTCAATTATGTCCAAATGGCTCTCATGAAAGTGGCGCGCATGACA 2209  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 2210 GGTACGTAGTTTATGACCATCTCACCCACTCTCGGAGCTGGGCCACCGGCGCTACGA 2269  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 2270 GACCTTGGCGTGGCAGTTGAGCCCGTCTTCTCTGATATGAGAGACCAAGGTTATCACC 2329  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 2330 TGGGGGAGAGACACCGCGCGCTGTGGGGACATCATCTTGGGCTGCGCTCTCCGCCCGC 2389  
 Db |||||  
 QY 1804 ----- 1803  
 Db |||||  
 QY 2390 AGGGGAGGAGATACATCTGGGACCGGACAGACCTTGAAGGGCAGGGGTGGCGACTC 2449  
 Db |||||  
 QY 1804 ---GGGCTTATTACGGCTTACTCCCAACAGACGCGAGGCTTCTTGGCTGCAATCATCT 1860  
 Db |||||  
 QY 2450 CTCGCGCTTATTACGGCTTACTCCCAACAGACGCGAGGCTTCTTGGCTGCAATCATCT 2509  
 Db |||||  
 QY 1861 AGCCTCACAGGCGGACAGGAAACAGGTCGAGGGGAGTCCAGTGGTCTCCACCGCA 1920  
 Db |||||  
 QY 2510 AGCCTCACAGGCGGACAGGAAACAGGTCGAGGGGAGTCCAGTGGTCTCCACCGCA 2569  
 Db |||||  
 QY 1921 ACACATCTTCTCTCGGACCTGCGTCAATGGCGTGTGTGGACTGTCTTATCATGTTGCC 1980  
 Db |||||  
 QY 2570 ACACATCTTCTCTCGGACCTGCGTCAATGGCGTGTGTGGACTGTCTTATCATGTTGCC 2629  
 Db |||||  
 QY 1981 GGCTCAAGACCTTCTCGCGCGCAAGGCGCCATCACCAANTGTACCAANTGTGGAC 2040  
 Db |||||

Db 2630 GGCTCAAAGACCCTTGCGGCCCAAGAGGCCCAATCACCCAAATGTACACCAATGTGGAC 2689  
QY 2041 CAGGACCTCGTGGCTGGGAGCGCCCGCGGGCGGCTTCTTTGACACCAATGCACTGCG 2100  
Db 2690 CAGGACCTCGTGGCTGGGAGCGCCCGCGGGCGGCTTCTTTGACACCAATGCACTGCG 2749  
QY 2101 GGCAGCTCGGACCTTTACTTGGTTCACGAGCATGCGCATGTCATTTCCGGTGGCGCGCGG 2160  
Db 2750 GGCAGCTCGGACCTTTACTTGGTTCACGAGCATGCGCATGTCATTTCCGGTGGCGCGCGG 2809  
QY 2161 GGCAGCAGCGGGAGCCTACTCTCCCGCAGGCGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
Db 2810 GGCAGCAGCGGGAGCCTACTCTCCCGCAGGCGCGCTCTCTACTTTGAAGGGCTCTTCG 2869  
QY 2221 GCGGTTCACCTGCTGCGCCCTCGGGGACGCTGTGGGCACTTTCCGGGCTGCGGTGCG 2280  
Db 2870 GCGGTTCACCTGCTGCGCCCTCGGGGACGCTGTGGGCACTTTCCGGGCTGCGGTGCG 2929  
QY 2281 ACCGAGGGGTGGGAAGGGGTGCACTTTGTACCCGTCGAGTCTATGGAAACCACTATG 2340  
Db 2930 ACCGAGGGGTGGGAAGGGGTGGAATTTGTACCCGTCGAGTCTATGGAAACCACTATG 2989  
QY 2341 CGTCCCGGCTTTACGGCAACTCGTCCCTCCGCGCGGTACCGCAGACATTCAGGGT 2400  
Db 2990 CGTCCCGGCTTTACGGCAACTCGTCCCTCCGCGCGGTACCGCAGACATTCAGGGT 3049  
QY 2401 GCCATCTACGCCCTACTGTTAGGGGCAAGAGCACTAAGTGGCGGCTGCGTATGCA 2460  
Db 3050 GCCATCTACGCCCTACTGTTAGGGGCAAGAGCACTAAGTGGCGGCTGCGTATGCA 3109  
QY 2461 GCCAAGGGTATAGGTGCTGCTCAACCCGCTCGCGCGCACTAGGTTCGGG 2520  
Db 3110 GCCAAGGGTATAGGTGCTGCTCAACCCGCTCGCGCGCACTAGGTTCGGG 3169  
QY 2521 GCGTATATGTAAAGCACATGTTATCGACCTTAACATCAAGAACCGGGGTAAGGACCATC 2580  
Db 3170 GCGTATATGTAAAGCACATGTTATCGACCTTAACATCAAGAACCGGGGTAAGGACCATC 3229  
QY 2581 ACCAGGGTGGCCCATCAGTACTCACCTATGGCAAGTTTCTGCGGAGCGGTGTTGC 2640  
Db 3230 ACCAGGGTGGCCCATCAGTACTCACCTATGGCAAGTTTCTGCGGAGCGGTGTTGC 3289  
QY 2641 TCTGGGGCGCTATGACATCAATATGATGATGAGTGCCACTCAACTGACTCGACCAT 2700  
Db 3290 TCTGGGGCGCTATGACATCAATATGATGATGAGTGCCACTCAACTGACTCGACCAT 3349  
QY 2701 ATCTGGGCATCGGCACAGTCTTGACCAAGCGGAGACGCGTGAGCGCGCACTCGTCGTG 2760  
Db 3350 ATCTGGGCATCGGCACAGTCTTGACCAAGCGGAGACGCGTGAGCGCGCACTCGTCGTG 3409  
QY 2761 CTGCCACCGCTAGCGCTCGGGATCGGTACCGTGCCACATCCAAACATCGAGAGGTG 2820  
Db 3410 CTGCCACCGCTAGCGCTCGGGATCGGTACCGTGCCACATCCAAACATCGAGAGGTG 3469  
QY 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCAATCCCAATCGAGACCATC 2880  
Db 3470 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCAATCCCAATCGAGACCATC 3529  
QY 2881 AAGGGGGAGGACCTCACTTTCTGCGCATTCGAAGAAAGAAATGATGAGTCTGCGCGG 2940  
Db 3530 AAGGGGGAGGACCTCACTTTCTGCGCATTCGAAGAAAGAAATGATGAGTCTGCGCGG 3589  
QY 2941 AAGTGTCCGCGCTCGGACCTCAATGCTGTAGCATATTTACGGGGCTTGATGATCCGTC 3000  
Db 3590 AAGTGTCCGCGCTCGGACCTCAATGCTGTAGCATATTTACGGGGCTTGATGATCCGTC 3649  
QY 3001 ATACCAACTAGCGAGACGTCATTTGCTGAGCAACGAGCGCTCTAATGAGCGGGTTTACC 3060  
Db 3650 ATACCAACTAGCGAGACGTCATTTGCTGAGCAACGAGCGCTCTAATGAGCGGGTTTACC 3709  
QY 3061 GCGGATTTGCACTCAGTGATCGACTCGAATACATGTGTCAACCCAGACAGTCGACTTCAGC 3120  
Db 3710 GCGGATTTGCACTCAGTGATCGACTCGAATACATGTGTCAACCCAGACAGTCGACTTCAGC 3769

QY 3121 CTGGAACCGGACCTTTCACATTTAGACGACACCGTGCACAGACGCGGTGTGACGCTCG 3180  
Db 3770 CTGGAACCGGACCTTTCACATTTAGACGACACCGTGCACAGACGCGGTGTGACGCTCG 3829  
QY 3181 CAGCGGCGAGGACGACTGCTAGGGGAGATGGGCAATTTACAGTTTGTGACTCCAGGA 3240  
Db 3830 CAGCGGCGAGGACGACTGCTAGGGGAGATGGGCAATTTACAGTTTGTGACTCCAGGA 3889  
QY 3241 GAAACGGCCCTCGGGCAATGTTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGTGT 3300  
Db 3890 GAAACGGCCCTCGGGCAATGTTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGTGT 3949  
QY 3301 GCTTGGTACAGACTCAGCCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTTAAACACA 3360  
Db 3950 GCTTGGTACAGACTCAGCCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTTAAACACA 4009  
QY 3361 CAGGGTGTGCGGCTCTGCCAGGACCATCTCGAGTTCTGGGAGAGCGTCTTTACAGGGCTC 3420  
Db 4010 CAGGGTGTGCGGCTCTGCCAGGACCATCTCGAGTTCTGGGAGAGCGTCTTTACAGGGCTC 4069  
QY 3421 ACCACATAGACGCCAATTTCTTGTCCAGACTAAGCAGCAGGAGACAACTTCCCTTAC 3480  
Db 4070 ACCACATAGACGCCAATTTCTTGTCCAGACTAAGCAGCAGGAGACAACTTCCCTTAC 4129  
QY 3481 CTGGTAGCATACGAGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCATCGTGGGAC 3540  
Db 4130 CTGGTAGCATACGAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCATCGTGGGAC 4189  
QY 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTACGCTGACAGGGGCAACGCGGCTCGT 3600  
Db 4190 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACAGGGGCAACGCGGCTCGT 4249  
QY 3601 TATAGGCTGGAGCCGTTTCAAAAAGAGTTTACTACCAACACCCCAATPAAACCAATACATC 3660  
Db 4250 TATAGGCTGGAGCCGTTTCAAAAAGAGTTTACTACCAACACCCCAATPAAACCAATACATC 4309  
QY 3661 ATGGCATGATGTGCGGTGACCTGGAGTGTGCTCAGAGACCTGGGTGCTGGTAGGGGA 3720  
Db 4310 ATGGCATGATGTGCGGTGACCTGGAGTGTGCTCAGAGACCTGGGTGCTGGTAGGGGA 4369  
QY 3721 GTCTAGACGCTTGCGCGGTATGCTGCAACAGGAGCGGTGTCATTTGTTGGGCGG 3780  
Db 4370 GTCTAGACGCTTGCGCGGTATGCTGCAACAGGAGCGGTGTCATTTGTTGGGCGG 4429  
QY 3781 ATCATCTTGTCCGGAAGCGCGCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 4430 ATCATCTTGTCCGGAAGCGCGCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 4489  
QY 3841 GATGAGATGGAAGTGGGCTCACACCTTCTTATCATCGAACAGGGAATGCACTGCGC 3900  
Db 4490 GATGAGATGGAAGTGGGCTCACACCTTCTTATCATCGAACAGGGAATGCACTGCGC 4549  
QY 3901 GAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGCT 3960  
Db 4550 GAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGCT 4609  
QY 3961 GCTGCTCCCGTGGTGAATCCAAAGTGGGACCGCTCGAAGCCCTTCTGGGCGAAGCATATG 4020  
Db 4610 GCTGCTCCCGTGGTGAATCCAAAGTGGGACCGCTCGAAGCCCTTCTGGGCGAAGCATATG 4669  
QY 4021 TGGAAATTCATAGCGGGATACAAATATTTAGAGGCTTGTGCACTGTGCTGCTGGAACCCC 4080  
Db 4670 TGGAAATTCATAGCGGGATACAAATATTTAGAGGCTTGTGCACTGTGCTGCTGGAACCCC 4729  
QY 4081 GCGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCCCGCTCACCCACCAACAT 4140  
Db 4730 GCGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCCCGCTCACCCACCAACAT 4789  
QY 4141 ACCCTCTGTTTAAACATCTCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGGCT 4200  
Db 4790 ACCCTCTGTTTAAACATCTCTGGGGGATGGGTGGCGGCCCAACTTGTCTCTCCAGGCT 4849





```
Db 7010 CTGGAAGACTAGACACCAATTCACACCAATATGCGCAAAAAATGAGTTTCTGCG 7069
QY 6421 GTCCAAACAGAGAGAGGGGGCGCAAGCCAGCTCCGCTTATCGTATTCCTCCAGATTGGGG 6480
Db 7070 GTCCAAACAGAGAGAGGGGGCGCAAGCCAGCTCCGCTTATCGTATTCCTCCAGATTGGGG 7129
QY 6481 GTTCGTGTGTGCGAGAAATGCCCCCTTTACGATGTGGTCTCCACCCTCCCTCAGGCCGTG 6540
Db 7130 GTTCGTGTGTGCGAGAAATGCCCCCTTTACGATGTGGTCTCCACCCTCCCTCAGGCCGTG 7189
QY 6541 ATGGGCTCTTATACGGAATCCAAATCTCTCTGACAGAGGGGTCCAGTTCCTGTTGAAT 6600
Db 7190 ATGGGCTCTTATACGGAATCCAAATCTCTCTGACAGAGGGGTCCAGTTCCTGTTGAAT 7249
QY 6601 GCCTGAAAGCGAAGAAATGCCCTATGGCTTCGCAATATGACACCGGCTGTTTGAATCA 6660
Db 7250 GCTTGGAAAGCGAAGAAATGCCCTATGGCTTCGCAATATGACACCGGCTGTTTGAATCA 7309
QY 6661 ACGGTCACTGAGAAATGACATCCGTGTTGAGGAGTCAATCTACCAATGTTGTGACTTGCC 6720
Db 7310 ACGGTCACTGAGAAATGACATCCGTGTTGAGGAGTCAATCTACCAATGTTGTGACTTGCC 7369
QY 6721 CCCGAGCAGACAGGCGCATAGGTCGTCTCAGAGCGGCTTTACATCGGGGGCCCCCTG 6780
Db 7370 CCCGAGCAGACAGGCGCATAGGTCGTCTCAGAGCGGCTTTACATCGGGGGCCCCCTG 7429
QY 6781 ACTAATCTAAAGGCGAAGCTGGGCTATCGCGGTGCGCGGCGAGCGGTGACTGACG 6840
Db 7430 ACTAATCTAAAGGCGAAGCTGGGCTATCGCGGTGCGCGGCGAGCGGTGACTGACG 7489
QY 6841 ACCAGTCGGTAAATPACCTCACAATGTTACTTGAAGCGCGTGCAGGCTGTCGAGCTGCG 6900
Db 7490 ACCAGTCGGTAAATPACCTCACAATGTTACTTGAAGCGCGTGCAGGCTGTCGAGCTGCG 7549
QY 6901 AAGCTCAGGACTGACGATGCTGATGCGGAGACCTTGTGCTTATCTGTAAGC 6960
Db 7550 AAGCTCAGGACTGACGATGCTGATGCGGAGACCTTGTGCTTATCTGTAAGC 7609
QY 6961 GCGGGACCCAGAGAGAGGCGAGCCTAGCGGCTTACGGAGGCTATGACTAGATAC 7020
Db 7610 GCGGGACCCAGAGAGAGGCGAGCCTAGCGGCTTACGGAGGCTATGACTAGATAC 7669
QY 7021 TCTGCCCCCTCGGGGACCCGCCAAACAGAAATACGACTTGGAGTTGATAAATCATGCG 7080
Db 7670 TCTGCCCCCTCGGGGACCCGCCAAACAGAAATACGACTTGGAGTTGATAAATCATGCG 7729
QY 7081 TCCTCCAAATGTCAGTGGCGCAGATGCACTGCGCAAAAGGGTGTACTATCTCACCCGT 7140
Db 7730 TCCTCCAAATGTCAGTGGCGCAGATGCACTGCGCAAAAGGGTGTACTATCTCACCCGT 7789
QY 7141 GACCCACCCACCCCTTGGCGGGCTGCGTGGGAGACAGTACACACTCCAGTCAAT 7200
Db 7790 GACCCACCCACCCCTTGGCGGGCTGCGTGGGAGACAGTACACACTCCAGTCAAT 7849
QY 7201 TCCTGGGTAGGCAACATCATGATATGCGGCCACCTTGTGGCAAGGATCATCTGATG 7260
Db 7850 TCCTGGGTAGGCAACATCATGATATGCGGCCACCTTGTGGCAAGGATCATCTGATG 7909
QY 7261 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAAACAACTTGAAAGAGCCCTAGATTGTGAG 7320
Db 7910 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAAACAACTTGAAAGAGCCCTAGATTGTGAG 7969
QY 7321 ATCTACGGGGCTGTATCTCATTTGAGCACTTACCTTACCTCAGATCATTCACGACTC 7380
Db 7970 ATCTACGGGGCTGTATCTCATTTGAGCACTTACCTTACCTCAGATCATTCACGACTC 8029
QY 7381 CATGGCTTATGCGCATTTTCACTCCATAGTTACTTCTCCAGTGTAGATCAATAGGGTGGCT 7440
Db 8030 CACGGCTTATGCGCATTTTCACTCCATAGTTACTTCTCCAGTGTAGATCAATAGGGTGGCT 8089
QY 7441 TCATGCTTCAGAAACTTGGGGTACCAGCTTGGAGTCTGGAGACATCGGGCCAGAAAGT 7500
Db 8090 TCATGCTTCAGAAACTTGGGGTACCAGCTTGGAGTCTGGAGACATCGGGCCAGAAAGT 8149
```

```
QY 7501 GTCCGCGCTAGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560
Db 8150 GTCCGCGCTAGCTACTGTCCAGGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 8209
QY 7561 AACTGGGCGAGTAAGAGCAACAACTCAAACTCACTCCCAATCCCGGCTGCGTCCAGTTGGAT 7620
Db 8210 AACTGGGCGAGTAAGAGCAACAACTCAAACTCACTCCCAATCCCGGCTGCGTCCAGTTGGAT 8269
QY 7621 TTATCCAGCTGGTTCGTTGCTGTGTGTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 7680
Db 8270 TTATCCAGCTGGTTCGTTGCTGTGTGTACAGCGGGGAGACATATATCACAGCCTGTCTCGT 8329
QY 7681 GCCCGACCCCGCTGGTTCATGTGTGCTACTCTCTACTCTTCTGTAGGGGTAGGCACTAT 7740
Db 8330 GCCCGACCCCGCTGGTTCATGTGTGCTACTCTCTACTCTTCTGTAGGGGTAGGCACTAT 8389
QY 7741 CTACTCCCCCAACCGGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTGTTTT 7800
Db 8390 CTACTCCCCCAACCGGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTTT 8449
QY 7801 TTTCCC-----TTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7850
Db 8450 TTTCCCCTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 8509
QY 7851 TTCTCCTTTTTTTTTTCTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7910
Db 8510 CCCCCCTTTTTTTTCCCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 8569
QY 7911 TAGTCACGGCTAGCTGTGAAAGGTCCTGAGCGCTTGACCTGACAGAGAGTCTGATACGTG 7970
Db 8570 TAGTCACGGCTAGCTGTGAAAGGTCCTGAGCGCTTGACCTGACAGAGAGTCTGATACGTG 7970
QY 7971 GCCTCTCTGCAAGATCAAGT 7989
Db 8630 GCCTCTCTGCAAGATCAAGT 8648
```

RESULT 15  
US-10-029-907-4  
; Sequence 4, Application US/10029907  
; Patent No. 6706874  
; GENERAL INFORMATION:  
; APPLICANT: BOEHRINGER INGELHEIM (CANADA) LTD.  
; TITLE OF INVENTION: SELF REPLICATING RNA MOLECULE FROM  
; TITLE OF INVENTION: HEPATITIS C VIRUS  
; FILE REFERENCE: 13/083  
; CURRENT APPLICATION NUMBER: US/10/029,907  
; PRIOR FILING DATE: 2001-12-21  
; PRIOR APPLICATION NUMBER: 60/257,857  
; PRIOR FILING DATE: 2000-12-22  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 8643  
; TYPE: DNA  
; ORGANISM: HCV  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1802)...(8407)  
US-10-029-907-4

Query Match 90.4%; Score 7221; DB 4; Length 8643;  
Best Local Similarity 91.9%; Pred. No. 0;  
Matches 7952; Conservative 0; Mismatches 25; Indels 676; Gaps 4;

```
QY 2 CCAGCCCCCATTTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTCTGT 61
Db 2 CCAGCCCCCATTTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTCTGT 61
QY 62 CTTTCAGCAGAAAGCGTCTAGCCATGCGTTAGTATGAGTGTGTCGAGCCTCCAGGACC 121
Db 62 CTTTCAGCAGAAAGCGTCTAGCCATGCGTTAGTATGAGTGTGTCGAGCCTCCAGGACC 121
```

Qy	122	CCCCCTCCGGAGACCATAGTGTCTGGGAAACGGTGAATACACCGGAATTGCCAGG	181	Qy	1190	AGACCAACACGGTTCCCTCTAGCGGATCAATTCGCGCCCTCTCCCTCCCCCCCCCTA	1249
Db	122	CCCCCTCCGGAGAGACCATAGTGTCTGGGAAACGGTGAATACACCGGAATTGCCAGG	181	Db	1202	AGACCAACACGGTTCCCTCTAGCGGATCAATTCG-----CCCCCCCCCTA	1250
Qy	182	ACGACCGGTCCTTTCTTGATCAACCCGCTCAATGCTGGAGATTGGGCGTGCCTCCG	241	Qy	1250	AGTTACTGGCGAAGCCGCTTGGAAATAAGGCGGTGTGGTTTGTCTATATATTTT	1309
Db	182	ACGACCGGTCCTTTCTTGATCAACCCGCTCAATGCTGGAGATTGGGCGTGCCTCCG	241	Db	1251	ACGTTACTGGCGAAGCCGCTTGGAAATAAGGCGGTGTGGTTTGTCTATATATTTT	1310
Qy	242	CGAGATCTGTAGCCGATGATGTGGGTGCGAAAGGCTTGTGTACTGCTGTATAGG	301	Qy	1310	CCACCATATTGCGCTCTTTTGGCAATGTGAGGCGCCGGAACCTGCGCCCTGTCTCTTGA	1369
Db	242	CGAGATCTGTAGCCGATGATGTGGGTGCGAAAGGCTTGTGTACTGCTGTATAGG	301	Db	1311	CCACCATATTGCGCTCTTTTGGCAATGTGAGGCGCCGGAACCTGCGCCCTGTCTCTTGA	1370
Qy	302	TGCTTCGGATGCCCCGGAGGTCCTGTAGACCGTGCACCATGAGCAGCAATCCTAAAC	361	Qy	1370	CGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGAATGTCTG	1429
Db	302	TGCTTCGGATGCCCCGGAGGTCCTGTAGACCGTGCACCATGAGCAGCAATCCTAAAC	361	Db	1371	CGAGCATTCCTAGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTGAATGTCTG	1430
Qy	362	TCAAGAAAAACCAAGGGCGCCCATGATTGAACAAGATGATGACGACGATCTCTCC	421	Qy	1430	TGAAGAAAGAGTTCTCTGGAAGCTTCTTGAAGCAAAACAAGCTGTAGACACCTTT	1489
Db	362	TCAAGAAAAACCAAGGGCGCCCATGATTGAACAAGATGATGACGACGATCTCTCC	421	Db	1431	TGAAGAAAGAGTTCTCTGGAAGCTTCTTGAAGCAAAACAAGCTGTAGACACCTTT	1490
Qy	422	GGCGCTTGGGTGGAGGCTATTCCGCTATGATGCTGGGCAACACACATCCGCTGCTC	481	Qy	1490	GCAGGACGGAACCCCACTTGGCGACAGGTGCTCTGCGGCCAAAAGCCACGCTGAT	1549
Db	422	GGCGCTTGGGTGGAGGCTATTCCGCTATGATGCTGGGCAACACACATCCGCTGCTC	481	Db	1491	GCAGGACGGAACCCCACTTGGCGACAGGTGCTCTGCGGCCAAAAGCCACGCTGAT	1550
Qy	482	TGATGCGCGCTGCTCCGCTGTACGCGAGGGCGCCCGGTTCTTTTGTCAAGACCGA	541	Qy	1550	AAGATACCTGCAAAAGGCGGCAACCCCACTGCGGCTGCGGCTGTGATAGTTGTGG	1609
Db	482	TGATGCGCGCTGCTCCGCTGTACGCGAGGGCGCCCGGTTCTTTTGTCAAGACCGA	541	Db	1551	AAGATACCTGCAAAAGGCGGCAACCCCACTGCGGCTGCGGCTGTGATAGTTGTGG	1610
Qy	542	CTGTCCGCTGCTCCTGAATGAATGCAAGGAGGCGCGGCTATCTGTGGTGGCCAC	601	Qy	1610	AAAGGTCAAAATGGCTCTCTGAGCGTATTCAAAAGGGGCTCAAGGATGCCAGG	1669
Db	542	CTGTCCGCTGCTCCTGAATGAATGCAAGGAGGCGCGGCTATCTGTGGTGGCCAC	601	Db	1611	AAAGGTCAAAATGGCTCTCTGAGCGTATTCAAAAGGGGCTCAAGGATGCCAGG	1670
Qy	602	GACGGGCTTCTTGGCGAGCTGTGCTGACGTTGTCACTGAAGCGGAAGGACTGGCT	661	Qy	1670	TACCCATTGTATGGGATCTGATCTGGGCTCGGTGCAATGCTTTTACATGTTTGTAGT	1729
Db	602	GACGGGCTTCTTGGCGAGCTGTGCTGACGTTGTCACTGAAGCGGAAGGACTGGCT	661	Db	1671	TACCCATTGTATGGGATCTGATCTGGGCTCGGTGCAATGCTTTTACATGTTTGTAGT	1730
Qy	662	GCTATTGGGCGAGTGCAGGGGAGGATCTCTGTGATCTACCTTCTCTCCCGGAA	721	Qy	1730	CGAGGTTAAAAAAGCTCTAGGCCCCCGAAACCAAGGGGACGTTGTTCTTTGAAAAAC	1789
Db	662	GCTATTGGGCGAGTGCAGGGGAGGATCTCTGTGATCTACCTTCTCTCCCGGAA	721	Db	1731	CGAGGTTAGAAAAGCTCTAGGCCCCCGAAACCAAGGGGACGTTGTTCTTTGAAAAAC	1790
Qy	722	AGTATCCATCATGCTGATCAATGCGGCGGCTGCACTGCTGATCCGGCTACCTGCC	781	Qy	1790	ACGATTAATACCATG-----1803	1803
Db	722	AGTATCCATCATGCTGATCAATGCGGCGGCTGCACTGCTGATCCGGCTACCTGCC	781	Db	1791	ACGATTAATACCATGAGACCGGAGATGCGAGATCGTGGGAGGCGGCTTTTGTAGTCT	1850
Qy	782	ATTGACACCAAGCGAAACATCGATCGAGCGAGCACTGCTGATGGAAGCGGCTCT	841	Qy	1804	-----1803	1803
Db	782	ATTGACACCAAGCGAAACATCGATCGAGCGAGCACTGCTGATGGAAGCGGCTCT	841	Db	1851	TGATACTCTTGTCTACCGCATTAAGCTGTTCTCTGCTAGGCTCATATGTTGTT	1910
Qy	842	TGTCGATCAGATGATCTGGAAGAAGCATCAGGGGCTCGCGCCAGCGAACTGTTCCG	901	Qy	1804	-----1803	1803
Db	842	TGTCGATCAGATGATCTGGAAGAAGCATCAGGGGCTCGCGCCAGCGAACTGTTCCG	901	Db	1911	TACAATATTTATACACGAGGCGAGGACACATTGCAAGTGTGGATCCCCCCTCAACG	1970
Qy	902	CAGGCTCAAGCGCGCATGCCCCGACGCGAGGATCTGCTGTCGACCATGCGATGCTG	961	Qy	1804	-----1803	1803
Db	902	CAGGCTCAAGCGCGCATGCCCCGACGCGAGGATCTGCTGTCGACCATGCGATGCTG	961	Db	1971	TTCCGGGGGCGCGGATGCGGCTCATCTCTCAGCTGCGGATCCACCCAGAGCTAATCT	2030
Qy	962	CTTGCGGAATATCATGTTGGAATAATGCGGCTTTTCTGGAATTCATGAGTGTGGCGGCT	1021	Qy	1804	-----1803	1803
Db	962	CTTGCGGAATATCATGTTGGAATAATGCGGCTTTTCTGGAATTCATGAGTGTGGCGGCT	1021	Db	2031	TTACCATCAACAAAATCTTCTCGCCATCTCGTCCACTCATGTTGTTGCTCCAGGCTGTA	2090
Qy	1022	GGGTGTGGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATGCTGAAGACT	1081	Qy	1804	-----1803	1803
Db	1022	GGGTGTGGGACCGCTATCAGGACATAGCGTTGGCTACCGCTGATATGCTGAAGACT	1081	Db	2091	TAACCAAGTGGCTACTTCTGTCGCGCACACGCGGCTCATTCGTGATGATGCTGTTGTC	2150
Qy	1082	TGGCGGGAATGGCTGACCGCTTCTGCTTTTACGGTATCGCGCTCCCGATTCGCA	1141	Qy	1804	-----1803	1803
Db	1082	TGGCGGGAATGGCTGACCGCTTCTGCTTTTACGGTATCGCGCTCCCGATTCGCA	1141	Db	2151	GGAAGTGTCTGGGGTCAATATGTCAAAATGCTCTCATGAAGTTGGCGGCTGACAG	2210
Qy	1142	GCGATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGT-----TAAAC	1189	Qy	1804	-----1803	1803
Db	1142	GCGATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGT-----TAAAC	1201	Db	2211	GTAGTACGTTTATGACCATCTCACCCACTGCGGGACTGGGCCACGCGGCTTACGAG	2270
				Qy	1804	-----1803	1803

Db	2271	ACCTTGGGTGGCAGTTGAGCCCGTCGTCTTCTCTGATATGAGACCAAGGTTATCACT	2330
QY	1804	-----	1803
Db	2331	GGGGGGACACACCGCGCGTGTGGGGACATCATCTTGGGCCCTGCCCGTCTCCGCCCGCA	2390
QY	1804	-----	1803
Db	2391	GGGGAGGAGATACATCTGGGACCGGACAGACGCTTGAAGGGCAGGGGTGGGACTCC	2450
QY	1804	--GGCCCTATTACGCCCTACTCCCAACAGACGCGAGGCCCTACTTGGCTGCATCATCACTA	1861
Db	2451	TCGCGCCTATTAGCGCCTACTCCCAACAGACGCGAGGCCCTACTTGGCTGCATCATCACTA	2510
QY	1862	GCCTCAGAGCGGGACAGGAACAGAGTCGAGGGGGAGGTCCAAATGCTCTCCACCGAA	1921
Db	2511	GCCTCAGAGCGGGACAGGAACAGAGTCGAGGGGGAGGTCCAAATGCTCTCCACCGAA	2570
QY	1922	CACAACTTTCTGCGGACCTGCGTCAATGGCGTGTGTGGACTGTCTATCATGTGCGG	1981
Db	2571	CACAACTTTCTGCGGACCTGCGTCAATGGCGTGTGTGGACTGTCTATCATGTGCGG	2630
QY	1982	GCTCAAGACCTTTGCCGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGACC	2041
Db	2631	GCTCAAGACCTTTGCCGGCCCAAGGGCCCAATCACCCAAATGTACCAATGTGGACC	2690
QY	2042	AGGACCTCGTGGCTGGCAAGGCCGCCCGGGGCGCTTCTTGACACATGTGCACTGCG	2101
Db	2691	AGGACCTCGTGGCTGGCAAGGCCGCCCGGGGCGCTTCTTGACACATGTGCACTGCG	2750
QY	2102	GCAGCTCGGACCTTTACTTTGTCACGAGGCATGCCGATGTCATTCGGTGCAGCGCGG	2161
Db	2751	GCAGCTCGGACCTTTACTTTGTCACGAGGCATGCCGATGTCATTCGGTGCAGCGCGG	2810
QY	2162	GCACAGCAGGGGAGCCTACTCTCCCGCAGCCCGTCTCCTACTTGAAGGCTCTTCGG	2221
Db	2811	GCACAGCAGGGGAGCCTACTCTCCCGCAGCCCGTCTCCTACTTGAAGGCTCTTCGG	2870
QY	2222	CGGTCACATGCTCTGCCCTCGGGGACGCTGTGGGCATCTTTCGGGCTGCGGTGTGCA	2281
Db	2871	CGGTCACATGCTCTGCCCTCGGGGACGCTGTGGGCATCTTTCGGGCTGCGGTGTGCA	2930
QY	2282	CCCAGGGGTTGCAAGCGGTGACCTTTGACCGTCGAGTCTATGGAACCACTATGC	2341
Db	2931	CCCAGGGGTTGCAAGCGGTGACCTTTGACCGTCGAGTCTATGGAACCACTATGC	2990
QY	2342	GGTCCCGGCTTTCACGGACAACCTCGTCCCTCGCGCGGTACCGCAGACATTCAGGTGG	2401
Db	2991	GGTCCCGGCTTTCACGGACAACCTCGTCCCTCGCGCGGTACCGCAGACATTCAGGTGG	3050
QY	2402	CCCATCTACAGCCCTACTGCTAGCGGCAAGAGACTAAGGTGCGCGCTCGGTATGAG	2461
Db	3051	CCCATCTACAGCCCTACTGCTAGCGGCAAGAGACTAAGGTGCGCGCTCGGTATGAG	3110
QY	2462	CCCAGGGTATAAGGTGCTTCTGCTGAACCGCTCGCTCGCGCGCACCTAGGTTTCGGG	2521
Db	3111	CCCAGGGTATAAGGTGCTTCTGTAACCCGCTCGCTCGCGCGCACCTAGGTTTCGGG	3170
QY	2522	CGTATATGCTAAGGCAATGGTATCGACCCCTAATACAGAACCGGGGTAAAGCACTCA	2581
Db	3171	CGTATATGCTAAGGCAATGGTATCGACCCCTAATACAGAACCGGGGTAAAGCACTCA	3230
QY	2582	CCAGGGTGCCCCATCAGCTACTCCACCTATGGCAAGTTCTTTCGCGACGGTGGTGTCT	2641
Db	3231	CCAGGGTGCCCCATCAGCTACTCCACCTATGGCAAGTTCTTTCGCGACGGTGGTGTCT	3290
QY	2642	CTGGGGCGCTATGACATATAATATGTATGTAGTGCACCTCAACTGACTCGACCACTA	2701
Db	3291	CTGGGGCGCTATGACATATAATATGTATGTAGTGCACCTCAACTGACTCGACCACTA	3350
QY	2702	TCCTGGGCATCGGCACAGTCTCGGACCAAGCGGAGACGGCTGGAGCGGACTCGTCTGTC	2761

Db	3351	TCCTGGGCATCGGCACAGTCTCTGGACCAAGCGGACCGGCTGGAGCGGACTCGTCTGTC	3410
QY	2762	TCGCCACCGCTACCGCTCCGGGATCGGTCAACGTCGCCATCCAAACATCGAGGAGGTGG	2821
Db	3411	TCGCCACCGCTACCGCTCCGGGATCGGTCAACGTCGCCATCCAAACATCGAGGAGGTGG	3470
QY	2822	CTCTGTCCAGCACTGGAGAAATCCCTTTTATATGGCAAGCCATCCCATCGAGACCATCA	2881
Db	3471	CTCTGTCCAGCACTGGAGAAATCCCTTTTATATGGCAAGCCATCCCATCGAGACCATCA	3530
QY	2882	AGGGGGGAGGACACCTCATTTTCTGCCATTCACAAAGAAATGTGATGAGCTCGCGCGCA	2941
Db	3531	AGGGGGGAGGACACCTCATTTTCTGCCATTCACAAAGAAATGTGATGAGCTCGCGCGCA	3590
QY	2942	AGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTCACGGGGCCTTGTATATCCGTCA	3001
Db	3591	AGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTCACGGGGCCTTGTATATCCGTCA	3650
QY	3002	TACCAACTAGCGGAGACGCTCATTTGCTGTAGCAACGGACGCTCTAATGACGGGCTTTACG	3061
Db	3651	TACCAACTAGCGGAGACGCTCATTTGCTGTAGCAACGGACGCTCTAATGACGGGCTTTACG	3710
QY	3062	CGCATTTGCACTCAGTGTGATCGACTGCAATACATGTGTCAACCGACAGCTCGACTTCAG	3121
Db	3711	CGCATTTGCACTCAGTGTGATCGACTGCAATACATGTGTCAACCGACAGCTCGACTTCAG	3770
QY	3122	TGGACCCGACCTTCAACATTCAGACGACGACCGTGTGCCAAGACGCGGTGTACGCTCGC	3181
Db	3771	TGGACCCGACCTTCAACATTCAGACGACGACCGTGTGCCAAGACGCGGTGTACGCTCGC	3830
QY	3182	AGCGCGGAGGACGACTGTTAGGGGACGATGGGCACTTACAGGTTTGTGATCTCAGAG	3241
Db	3831	AGCGCGGAGGACGACTGTTAGGGGACGATGGGCACTTACAGGTTTGTGATCTCAGAG	3890
QY	3242	AACGGCCCTCGGGCATGTTTCGATTCCTCGGTTCTGTGAGTGTATCAGCGGCGCTGTG	3301
Db	3891	AACGGCCCTCGGGCATGTTTCGATTCCTCGGTTCTGTGAGTGTATCAGCGGCGCTGTG	3950
QY	3302	CTTGGTAGAGCTCAACCGCCCGCAGACCTCAGTTAGGTTTCGGGCTTACCTAAACACAC	3361
Db	3951	CTTGGTAGAGCTCAACCGCCCGCAGACCTCAGTTAGGTTTCGGGCTTACCTAAACACAC	4010
QY	3362	CAGGTTTCGGCTGTCCAGACCATCTGAGGTTCTGGGAGAGGCTCTTTACAGGCTCA	3421
Db	4011	CAGGTTTCGGCTGTCCAGACCATCTGAGGTTCTGGGAGAGGCTCTTTACAGGCTCA	4070
QY	3422	CCACATAGAGCGCCCATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTACC	3481
Db	4071	CCACATAGAGCGCCCATTTCTTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTTACC	4130
QY	3482	TGGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCTGGGGACC	3541
Db	4131	TGGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCTGGGGACC	4190
QY	3542	AAATGTGGAGTGTCTCATACGGCTAAGCCTACGCTGACCGGCGCAACGCCCTGTGT	3601
Db	4191	AAATGTGGAGTGTCTCATACGGCTAAGCCTACGCTGACCGGCGCAACGCCCTGTGT	4250
QY	3602	ATAGGCTGGAGCGGCTTCAAAACAGAGTTTACTACACACACCCCATCAACCAATACATCA	3661
Db	4251	ATAGGCTGGAGCGGCTTCAAAACAGAGTTTACTACACACACCCCATCAACCAATACATCA	4310
QY	3662	TGGCATGCATGTGGCTGACCTCGTAGGTCGTTCAGGACACCTGGGTGCTGGTAGCGGAG	3721
Db	4311	TGGCATGCATGTGGCTGACCTCGTAGGTCGTTCAGGACACCTGGGTGCTGGTAGCGGAG	4370
QY	3722	TCCTAGCAGCTCTGGCGCGCTATTCCTGACACACAGGACGAGGTCATTTGTGGGCGAG	3781
Db	4371	TCCTAGCAGCTCTGGCGCGCTATTCCTGACACACAGGACGAGGTCATTTGTGGGCGAG	4430
QY	3782	TCATTTTGTCCGAAAGCGCGCATCATTCGCCACAGGAGTCTTTTACGGGAGTTTCG	3841
Db	4431	TCATTTTGTCCGAAAGCGCGCATCATTCGCCACAGGAGTCTTTTACGGGAGTTTCG	4490

QY	3842	ATGAGTGAAGAGTCCGCTC	4981	CCACGGGCCCCGTCACG
DB	4491	ATGAGTGAAGAGTCCGCTC	5571	CCACGGGCCCCGTCACG
QY	3902	AACAATTCACACAGAGG	4982	CTGCTGAGGAGTACGTG
DB	4551	AACAATTCACACAGAGG	5631	CTGCTGAGGAGTACGTG
QY	3962	CTGCTCCCTGCTGGAAT	5042	CCACTACACAACTAAGT
DB	4611	CTGCTCCCTGCTGGAAT	5691	CCACTACACAACTAAGT
QY	4022	GGAAATTCATCAGCGG	5102	ATGGGCTGCGGTTCAC
DB	4671	GGAAATTCATCAGCGG	5810	ATGGGCTGCGGTTCAC
QY	4082	CGATAGCATCACTGAT	5162	CATTCTCTGCTCGGCT
DB	4731	CGATAGCATCACTGAT	5811	CATTCTCTGCTCGGCT
QY	4142	CCCTCTCTGTTAACTC	5222	CGGACGTAGAGTGTCT
DB	4791	CCCTCTCTGTTAACTC	5930	TGGACGTAGAGTGTCT
QY	4202	CTTCTGCTTTCTGAGG	5341	CTAAGCGTAGGCTGGC
DB	4851	CTTCTGCTTTCTGAGG	5990	CTAAGCGTAGGCTGGC
QY	4262	AGTGCTCTGTTGGAAT	5401	TGFTCTCGGCTTCTTT
DB	4911	AGTGCTCTGTTGGAAT	6050	TGFTCTCGGCTTCTTT
QY	4322	TAAAGTTCATGAGCGG	5461	TCATTCAGGCGCACTC
DB	4971	TAAAGTTCATGAGCGG	6110	TCATTCAGGCGCACTC
QY	4382	TCCTCTCCCTCGGCT	5521	CAGAAATAAGGTAGTA
DB	5031	TCCTCTCCCTCGGCT	6170	CAGAAATAAGGTAGTA
QY	4442	TGGGCCCCAGGGAGG	5581	GGGAAATATCCGTTCC
DB	5091	TGGGCCCCAGGGAGG	6230	GGGAAATATCCGTTCC
QY	4502	GTAACCAACGCTCTCC	5641	CCATATGGGACACCCG
DB	5151	GTAACCAACGCTCTCC	6290	CCATATGGGACACCCG
QY	4562	AGATCTCTCTAGTCTT	5701	ACGTCTCTCCAGTGTG
DB	5211	AGATCTCTCTAGTCTT	6350	ACGTCTCTCCAGTGTG
QY	4622	AGGACTCTCTCAGCC	5761	CTCCAAGGAGAGAGAG
DB	5271	AGGACTCTCTCAGCC	6410	CTCCAAGGAGAGAGAG
QY	4682	CGGTGTTGA CTGATTT	5821	AGCTCCGCCACABAGA
DB	5331	CGGTGTTGA CTGATTT	6470	AGCTCCGCCACABAGA
QY	4742	TCCCTCTTCTTCTATG	5881	CGGCTCTCTCTGACCA
DB	5391	TCCCTCTTCTTCTATG	6530	CGGCTCTCTCTGACCA
QY	4802	AAACCACTGCCCCTG	5941	CTTCCATGCCCCCTTT
DB	5451	AAACCACTGCCCCTG	6590	CTTCCATGCCCCCTTT
QY	4862	TCGTGGGCGCTTAGAC	6001	CTACCGTAAAGCGAGG
DB	5511	TCGTGGGCGCTTAGAC	6650	CTACCGTAAAGCGAGG
			6061	CAGGCGCCCTGATCAC

Db	6651	CGGGCGCCTGATCAGCCATGCGCTGCGGAGAAACCAAGCTGCCATCAATGACCTGA	6710
Qy	6062	GCAACTCTTTGCTCGTCAACAACCTTGGTCTATGCTAACAACTCTGCGAGCGAAGCC	6121
Db	6711	GCAACTCTTTGCTCGTCAACAACCTTGGTCTATGCTAACAACTCTGCGAGCGAAGCC	6770
Qy	6122	TGCGCAGAGAAGGTCAACCTTTGACAGACTGCGAGTCTGAGACCACTACCGGAGC	6181
Db	6771	TGCGCAGAGAAGGTCAACCTTTGACAGACTGCGAGTCTGAGACCACTACCGGAGC	6830
Qy	6182	TGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGGCTAAACTTCTATCCGTGGAGG	6241
Db	6831	TGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGGCTAAACTTCTATCCGTGGAGG	6890
Qy	6242	AAGCCTGTAAAGTGAAGCGGCGGCGGACATTCGCGCAGATCTAAATTTGGCTATGGGCAAGG	6301
Db	6891	AAGCCTGTAAAGTGAAGCGGCGGCGGACATTCGCGCAGATCTAAATTTGGCTATGGGCAAGG	6950
Qy	6302	ACGTCGGGAACCTATCCAGCAAGCGGTTAACACATCCGCTCCGTGTTGGAAGGACTTGC	6361
Db	6951	ACGTCGGGAACCTATCCAGCAAGCGGTTAACACATCCGCTCCGTGTTGGAAGGACTTGC	7010
Qy	6362	TGGAAGACACTGAGACACCAATTTGACACCAACCATCATGCGCAAAAAATGAGGTTTTCTGCG	6421
Db	7011	TGGAAGACACTGAGACACCAATTTGACACCAACCATCATGCGCAAAAAATGAGGTTTTCTGCG	7070
Qy	6422	TCCAAACAGAGAAGGGGGCGGCGGAGCGAGTTCGCTTATCGTATTTCCAGATTTGGGGG	6481
Db	7071	TCCAAACAGAGAAGGGGGCGGCGGAGCGAGTTCGCTTATCGTATTTCCAGATTTGGGGG	7130
Qy	6482	TTGCTGTGTCGAGAAAATGSCCTTTTACAGATGTTGCTCCACCTCCCTCAGGCGGTGA	6541
Db	7131	TTGCTGTGTCGAGAAAATGSCCTTTTACAGATGTTGCTCCACCTCCCTCAGGCGGTGA	7190
Qy	6542	TGGGCTCTTACAGGATTCGAATACTCTCTGAGACAGGGGTCGAGTTCTCTGTTGAATG	6601
Db	7191	TGGGCTCTTACAGGATTCGAATACTCTCTGAGACAGGGGTCGAGTTCTCTGTTGAATG	7250
Qy	6602	CCTGGAAGCGAAGAAATGCGCTATGCGCTTACAGAGCGGGTCCGAGTTCTCTGTTGAATG	6661
Db	7251	CCTGGAAGCGAAGAAATGCGCTATGCGCTTACAGAGCGGGTCCGAGTTCTCTGTTGAATG	7310
Qy	6662	CGGTCAGTGAAGTACATCCGCTGTTGAGGAGTCAATCTACCAATGTTGACTTGGGCC	6721
Db	7311	CGGTCAGTGAAGTACATCCGCTGTTGAGGAGTCAATCTACCAATGTTGACTTGGGCC	7370
Qy	6722	CCGAAGCCAGACAGGCCATAAGGTGCTCAGAGCGGCTTTACATCGGGGCGCCCTGA	6781
Db	7371	CCGAAGCCAGACAGGCCATAAGGTGCTCAGAGCGGCTTTACATCGGGGCGCCCTGA	7430
Qy	6782	CTAATTTAAAGGCGAGAACTGCGGCTATGCGCGGTGCGCGGAGCGGTGACTGACGA	6841
Db	7431	CTAATTTAAAGGCGAGAACTGCGGCTATGCGCGGTGCGCGGAGCGGTGACTGACGA	7490
Qy	6842	CCAGTGGGTAATACCTCACTGTTGAGGAGTCAATCTACCAATGTTGACTTGGGCC	6901
Db	7491	CCAGTGGGTAATACCTCACTGTTGAGGAGTCAATCTACCAATGTTGACTTGGGCC	7550
Qy	6902	AGCTCCAGGACTGACGATGCTCGTATGCGGAGACGACTTGTGTTATCTGTGAAGCG	6961
Db	7551	AGCTCCAGGACTGACGATGCTCGTATGCGGAGACGACTTGTGTTATCTGTGAAGCG	7610
Qy	6962	CGGGACCCAAAGAGACGAGCGGCGCTACCGGCTTCCAGGAGGCTATGACTAGATACT	7021
Db	7611	CGGGACCCAAAGAGACGAGCGGCGCTACCGGCTTCCAGGAGGCTATGACTAGATACT	7670
Qy	7022	CTGCCCCCTTGGGACCCGCGGCAACAGATACGACTTGGAGTTGATAATCATGCT	7081
Db	7671	CTGCCCCCTTGGGACCCGCGGCAACAGATACGACTTGGAGTTGATAATCATGCT	7730
Qy	7082	CCTCCAATGTGTGAGTCCGCGCAGCATGCTGCGCAAAAGGCTGACTATCTCACCCTG	7141

Search completed: December 9, 2004, 12:43:49  
Job time : 779 secs

Db	7731	CCTCCAATGTGTGACTGCGGCACGATGATCTGGCAAAAGGCTGACTATCTCACCCGTG	7790
Qy	7142	ACCCACACACCCCTTTCGGGGGCTGCGTGGAGACAGCTAGACACATCTCAGTCAAT	7201
Db	7791	ACCCACACACCCCTTTCGGGGGCTGCGTGGAGACAGCTAGACACATCTCAGTCAAT	7850
Qy	7202	CCTGGCTAGGCAACATCATGATGCGGCCACCTTGTGGCAAGGATGATCTTGATGA	7261
Db	7851	CCTGGCTAGGCAACATCATGATGCGGCCACCTTGTGGCAAGGATGATCTTGATGA	7910
Qy	7262	CTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACCTTGAAGCCCTTAGATTTGTCAGA	7321
Db	7911	CTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACCTTGAAGCCCTTAGATTTGTCAGA	7970
Qy	7322	TCTACGGGCTGTTPACTCCATTTAGCCACTTGACCTACCTCAGATCAATCAACGACTCC	7381
Db	7971	TCTACGGGCTGTTPACTCCATTTAGCCACTTGACCTACCTCAGATCAATCAACGACTCC	8030
Qy	7382	ATGGCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGTTGGCT	7441
Db	8031	ACGGCTTAGCGCATTTTCACTCCATAGTACTCTCCAGGTGAGATCAATAGGTTGGCT	8090
Qy	7442	CATGCTCAGGAAACTTGGGCTACCGCCCTTTCGAGTCTGGAGACATCGGCCAGAAAGT	7501
Db	8091	CATGCTCAGGAAACTTGGGCTACCGCCCTTTCGAGTCTGGAGACATCGGCCAGAAAGT	8150
Qy	7502	TCCGGCTTAGGCTACTGTCTCCAGGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTCA	7561
Db	8151	TCCGGCTTAGGCTACTGTCTCCAGGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTCA	8210
Qy	7562	ACTGGGCAAGTAAAGCAACGAGCTCAAACTCACTCCAAATCCCGGCTGCGTCCGAGTTG	7621
Db	8211	ACTGGGCAAGTAAAGCAACGAGCTCAAACTCACTCCAAATCCCGGCTGCGTCCGAGTTG	8270
Qy	7622	TATCCAGTGTGTTGCTGCTTACAGCGGGGAGACATATATCACAGCCTGCTCGTG	7681
Db	8271	TATCCAGTGTGTTGCTGCTTACAGCGGGGAGACATATATCACAGCCTGCTCGTG	8330
Qy	7682	CCGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	7741
Db	8331	CCGACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	8390
Qy	7742	TACTCCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAAATAGGCCATCTCTGTTT	7801
Db	8391	TACTCCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAAATAGGCCATCTCTGTTT	8450
Qy	7802	TTCCCTTT	7858
Db	8451	TTCCCTTT	8510
Qy	7859	TTTTTTTCTCTTT	7916
Db	8511	TTTTTTTCTCTTT	8570
Qy	7917	CGGCTAGCTGTAAGGTCCGAGCGCTTGAATGAGAGAGTCTGATGAGAGTCTGATGAGTCT	7976
Db	8571	CGGCTAGCTGTAAGGTCCGAGCGCTTGAATGAGAGAGTCTGATGAGAGTCTGATGAGTCT	8630
Qy	7977	CTGAGATCAAGT 7989	
Db	8631	CTGAGATCAAGT 8643	

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 20:37:05 ; Search time 3599 Seconds  
(without alignments)  
12238.821 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989

Sequence: 1 gccagcccccagatgggggc.....ggcctctctgcagatcaagt 7989

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4085731 seqs, 2756760397 residues

Total number of hits satisfying chosen parameters: 8171462

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA: \*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7987.4	100.0	7989	17	US-10-639-150-1
2	7987.4	100.0	7992	13	US-10-005-469-1
3	7987.4	100.0	7992	16	US-10-434-842-1
4	7987.4	100.0	10690	15	US-10-125-940-1
5	7987.4	100.0	10690	15	US-10-125-920-1
6	7987.4	100.0	10690	16	US-10-467-000-3
7	7985.8	100.0	7992	13	US-10-005-469-6
8	7985.8	100.0	7992	16	US-10-434-842-6
9	7982.6	99.9	7992	13	US-10-005-469-2
10	7982.6	99.9	7992	13	US-10-005-469-5
11	7982.6	99.9	7992	16	US-10-434-842-2
12	7982.6	99.9	7992	16	US-10-434-842-5

13	7982.6	99.9	7992	16	US-10-434-842-17	Sequence 17, Appl
14	7981	99.9	7992	13	US-10-005-469-4	Sequence 4, Appl
15	7981	99.9	7992	16	US-10-434-842-4	Sequence 4, Appl
16	7981	99.9	7992	16	US-10-434-842-15	Sequence 15, Appl
17	7971.4	99.8	7989	13	US-10-434-842-16	Sequence 16, Appl
18	7971.2	99.8	7995	13	US-10-005-469-3	Sequence 3, Appl
19	7971.2	99.8	7995	16	US-10-434-842-3	Sequence 2, Appl
20	7610.4	95.3	12305	18	US-10-422-323A-2	Sequence 1, Appl
21	7607.2	95.2	12315	18	US-10-422-323A-1	Sequence 1, Appl
22	7272.2	91.0	8639	13	US-10-029-907-1	Sequence 1, Appl
23	7272.2	91.0	8639	15	US-10-309-561-1	Sequence 1, Appl
24	7272.2	91.0	8639	17	US-10-789-355-1	Sequence 1, Appl
25	7272.2	91.0	8639	18	US-10-686-835-1	Sequence 1, Appl
26	7271.2	91.0	8638	13	US-10-029-907-24	Sequence 24, Appl
27	7271.2	91.0	8638	15	US-10-309-561-24	Sequence 24, Appl
28	7271.2	91.0	8638	17	US-10-789-355-24	Sequence 24, Appl
29	7271.2	91.0	8638	18	US-10-686-835-24	Sequence 24, Appl
30	7248.2	90.7	8638	13	US-10-029-907-7	Sequence 7, Appl
31	7248.2	90.7	8638	15	US-10-309-561-7	Sequence 7, Appl
32	7248.2	90.7	8638	17	US-10-789-355-7	Sequence 7, Appl
33	7248.2	90.7	8638	18	US-10-686-835-7	Sequence 7, Appl
34	7247.2	90.7	8638	13	US-10-029-907-25	Sequence 25, Appl
35	7247.2	90.7	8638	15	US-10-309-561-25	Sequence 25, Appl
36	7247.2	90.7	8638	17	US-10-789-355-25	Sequence 25, Appl
37	7247.2	90.7	8638	18	US-10-686-835-25	Sequence 25, Appl
38	7245.2	90.7	8642	13	US-10-029-907-2	Sequence 2, Appl
39	7245.2	90.7	8642	15	US-10-309-561-2	Sequence 2, Appl
40	7245.2	90.7	8642	17	US-10-789-355-2	Sequence 2, Appl
41	7245.2	90.7	8642	18	US-10-686-835-2	Sequence 2, Appl
42	7244	90.7	8638	13	US-10-029-907-6	Sequence 6, Appl
43	7244	90.7	8638	15	US-10-309-561-6	Sequence 6, Appl
44	7244	90.7	8638	17	US-10-789-355-6	Sequence 6, Appl
45	7244	90.7	8638	18	US-10-686-835-6	Sequence 6, Appl

#### ALIGNMENTS

RESULT 1  
US-10-639-150-1  
; Sequence 1, Application US/10639150  
; Publication No. US20040121975A1  
; GENERAL INFORMATION:  
; APPLICANT: BRISTOL-MYERS SQUIBB COMPANY  
; TITLE OF INVENTION: HEPATITIS C VIRUS ASSAYS  
; FILE REFERENCE: D0224 NP  
; CURRENT APPLICATION NUMBER: US/10/639,150  
; CURRENT FILING DATE: 2003-08-12  
; PRIOR APPLICATION NUMBER: US 60/402,661  
; PRIOR FILING DATE: 2002-08-12  
; NUMBER OF SEQ ID NOS: 5  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 1  
; LENGTH: 7989  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: HCV Replicon  
US-10-639-150-1

Query Match	100.0%	Score 7987.4;	DB 17;	Length 7989;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 7989;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	1	GCACGCCCCGATGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTCTG	60	
Db	1	GCACGCCCCGATGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTCTG	60	
QY	61	TCTTCACGAGAAACGGTCTAGCCATGCGGTAGTATGAGTGTGTCGTCAGCCTCCAGAC	120	
Db	61	TCTTCACGAGAAACGGTCTAGCCATGCGGTAGTATGAGTGTGTCGTCAGCCTCCAGAC	120	
QY	121	CCCCCTCCCGGAGAGCCATAGTGTCTGCGGAACCGGTGAGTACACCGGAATGCCAG	180	

121 CCCCCCTCCCGGAGAGCCATAGTGGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180  
 181 GAGACCGGGTCTCTTTCTTGATCAACCCCGCTCAATGCTTGAGATTTGGCGGTGCCCC 240  
 181 GAGACCGGGTCTCTTTCTTGATCAACCCCGCTCAATGCTTGAGATTTGGCGGTGCCCC 240  
 241 GCGAGACTGTAGCGAGTAGTCTGGGTGCGAAAGGCCCTTGTTGTTACTGCTGATAGG 300  
 241 GCGAGACTGTAGCGAGTAGTCTGGGTGCGAAAGGCCCTTGTTGTTACTGCTGATAGG 300  
 301 GTCTTTGCGAGTGCCTCCCGGAGGTCTCTGAGACCGGTGACCATGAGCACGAATCTTAAC 360  
 301 GTCTTTGCGAGTGCCTCCCGGAGGTCTCTGAGACCGGTGACCATGAGCACGAATCTTAAC 360  
 361 CTCAAGAAACAAAGAGGCGGCCCATGATTTGAAACAAAGATGATTTGCAACGAGTTCTC 420  
 361 CTCAAGAAACAAAGAGGCGGCCCATGATTTGAAACAAAGATGATTTGCAACGAGTTCTC 420  
 421 CGCCCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCACACAGCAATCGGTGCT 480  
 421 CGCCCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCACACAGCAATCGGTGCT 480  
 481 CTGATCCCGCTGTTCCCGCTGTACGCGCAGCGGCGCCCGGTCTTTTGTCAAGACCG 540  
 481 CTGATCCCGCTGTTCCCGCTGTACGCGCAGCGGCGCCCGGTCTTTTGTCAAGACCG 540  
 541 ACTGTCCGGTCCCTGATGAATGAACTGAGACAGGAGCGCGGTATCGTGTGCGCA 600  
 541 ACTGTCCGGTCCCTGATGAATGAACTGAGACAGGAGCGCGGTATCGTGTGCGCA 600  
 601 CGACGGCGCTTCTTCGCGAGCTGTCTGACCTGTGCTCATCTCATCTTGTCTCTGCCGAGA 720  
 601 CGACGGCGCTTCTTCGCGAGCTGTCTGACCTGTGCTCATCTCATCTTGTCTCTGCCGAGA 720  
 661 TGCTATTGGCGAAGTGCCGCGGCGAGATCTCTGTCTCATCTCATCTTGTCTCTGCCGAGA 720  
 721 AGTATCCATATGCGTGTATGCAATGCGGGGCTGATAGCTGTTGATTCGGCTACTGCC 780  
 721 AGTATCCATATGCGTGTATGCAATGCGGGGCTGATAGCTGTTGATTCGGCTACTGCC 780  
 781 CATTCACACCAAGCGAAACATCGCATCGAGCGACAGTACTCTCGATGGAAGCGGCTC 840  
 781 CATTCACACCAAGCGAAACATCGCATCGAGCGACAGTACTCTCGATGGAAGCGGCTC 840  
 841 TTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAATCTTCG 900  
 841 TTGTGATCAGGATGATCTGGACGAAGAGCATCAGGGGCTCGCGCCAGCCGAATCTTCG 900  
 901 CCAGGCTCAAGGGCGCATGCGCGAGCGGAGGATCTCGTGAACCATGCGGATGCGCT 960  
 901 CCAGGCTCAAGGGCGCATGCGCGAGCGGAGGATCTCGTGAACCATGCGGATGCGCT 960  
 961 GCTTCCGGAATCATGTGTGAAATGCGCGTCTTCTGATTCATCGACTGTGGCGGC 1020  
 961 GCTTCCGGAATCATGTGTGAAATGCGCGTCTTCTGATTCATCGACTGTGGCGGC 1020  
 1021 TGGGTGTGCGGACCGCTATCAGGACATAGCGTTGCTACCCGTGATATGCTGAAGAGC 1080  
 1021 TGGGTGTGCGGACCGCTATCAGGACATAGCGTTGCTACCCGTGATATGCTGAAGAGC 1080  
 1081 TTGGGGGCAATGGGCTGACCGCTCTCTGCTGTTACGGTATCGCGCTCCCGATTCG 1140  
 1081 TTGGGGGCAATGGGCTGACCGCTCTCTGCTGTTACGGTATCGCGCTCCCGATTCG 1140  
 1141 AGCGCATCGCTTCTATCGCTTCTTGAAGATTCCTGAGTTTAAACAGACCAACAG 1200  
 1141 AGCGCATCGCTTCTATCGCTTCTTGAAGATTCCTGAGTTTAAACAGACCAACAG 1200  
 1201 GTTTCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCTTAACGTTACTGGC 1260

1201 GTTTCCTCTAGCGGATCAATTCGCGCGCTCTCCCTCCCGCCCTTACGTTACTGGC 1260  
 1261 CGAAGCGCTTGGAATTAAGCGCGGTGTGGTTTGTCTATATGTTATTTTCCACATATTG 1320  
 1261 CGAAGCGCTTGGAATTAAGCGCGGTGTGGTTTGTCTATATGTTATTTTCCACATATTG 1320  
 1321 CGCTCTTTTGGGAATGTAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGATTCCT 1380  
 1321 CGCTCTTTTGGGAATGTAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGATTCCT 1380  
 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTCGTGAAGGAAGCA 1440  
 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTGTAATGTCGTGAAGGAAGCA 1440  
 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCTTTGAGGCGCGG 1500  
 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCTTTGAGGCGCGG 1500  
 1501 AACCCCGCACCTGGCGACAGGTGCTCTCGGCCAAAGCCACGTGTATAAGATACACCT 1560  
 1501 AACCCCGCACCTGGCGACAGGTGCTCTCGGCCAAAGCCACGTGTATAAGATACACCT 1560  
 1561 GCAAGGCGGCAACACCCAGTGCACCGTTGTGAGTTGGATAGTTTGGAAAAGAGTCAA 1620  
 1561 GCAAGGCGGCAACACCCAGTGCACCGTTGTGAGTTGGATAGTTTGGAAAAGAGTCAA 1620  
 1621 TGGCTCTCTCAAGCGTATTCAAAGGGGCTGAAGGATGCCAGAAAGGTACCCCATTTGT 1680  
 1621 TGGCTCTCTCAAGCGTATTCAAAGGGGCTGAAGGATGCCAGAAAGGTACCCCATTTGT 1680  
 1681 ATGGGATCTGATCTGGGCTCGGTGCATGCTTTACATGCTTTAGTTCGAGTTAAAA 1740  
 1681 ATGGGATCTGATCTGGGCTCGGTGCATGCTTTACATGCTTTAGTTCGAGTTAAAA 1740  
 1741 AACGTCTAGCGCCCGCCCAACACCGGAGCTGTGTTTCTTTTGAATAACACGATATACC 1800  
 1741 AACGTCTAGCGCCCGCCCAACACCGGAGCTGTGTTTCTTTTGAATAACACGATATACC 1800  
 1801 ATGGCGCTTATACGGCTTACTCCCAACAGACGCGAGGCGCTACTTGGTGTGATCATCACT 1860  
 1801 ATGGCGCTTATACGGCTTACTCCCAACAGACGCGAGGCGCTACTTGGTGTGATCATCACT 1860  
 1861 AGCTCACAGCGCGGACAGGACACAGGTCGAGGGGAGTCCAGTGTGCTCCACCGCA 1920  
 1861 AGCTCACAGCGCGGACAGGACACAGGTCGAGGGGAGTCCAGTGTGCTCCACCGCA 1920  
 1921 ACACATCTTCTGCGGACCTGCGCTCAATGGCGTGTGTTGAGCTGTCTATCATGTTGCC 1980  
 1921 ACACATCTTCTGCGGACCTGCGCTCAATGGCGTGTGTTGAGCTGTCTATCATGTTGCC 1980  
 1981 GGCTCAAGACCTTTCGCGCGCCAAAGGGCCCAATCACCAATGTACACCAATGTGGAC 2040  
 1981 GGCTCAAGACCTTTCGCGCGCCAAAGGGCCCAATCACCAATGTACACCAATGTGGAC 2040  
 2041 CAGGACTCGTGGCTGGCAAGCGCCCGCGGGCGCTTCTTGACACCATGCACCTGC 2100  
 2041 CAGGACTCGTGGCTGGCAAGCGCCCGCGGGCGCTTCTTGACACCATGCACCTGC 2100  
 2101 GGCAGCTCGGACCTTTACTTGGTTCAGAGGATGCGGATGTCATTCGCGTGGCGCGCGG 2160  
 2101 GGCAGCTCGGACCTTTACTTGGTTCAGAGGATGCGGATGTCATTCGCGTGGCGCGCGG 2160  
 2161 GCGGACAGGAGGGAGCGCTACTCTCCCGCAGCGCGCTCTCTTGAAGGCTCTTTCG 2220  
 2161 GCGGACAGGAGGGAGCGCTACTCTCCCGCAGCGCGCTCTCTTGAAGGCTCTTTCG 2220  
 2221 GCGGTCCTACTGCTCTGCGGCTCGGGGACCGCTGTGGGATCTTTCGCGTGGCGCTGTC 2280  
 2221 GCGGTCCTACTGCTCTGCGGCTCGGGGACCGCTGTGGGATCTTTCGCGTGGCGCTGTC 2280  
 2281 ACCGAGGGGTTCGGAAGCGGTGAGCTTTGTACCGTTCGAGTCTATGGAACCACTATG 2340  
 2281 ACCGAGGGGTTCGGAAGCGGTGAGCTTTGTACCGTTCGAGTCTATGGAACCACTATG 2340

QY 2341 CGGTCCCGGCTTTCACGGCAAACTCGTCCCTCGGCGGTACCGCAGACATTCAGGTG 2400  
DB |||||  
QY 2341 CGGTCCCGGCTTTCACGGCAAACTCGTCCCTCGGCGGTACCGCAGACATTCAGGTG 2400  
DB |||||  
QY 2401 GCCCATCTACACGCCCTTACTGGTAGCGCAAGAGCACTAAGGTGCGCGCTGCGTATGCA 2460  
DB |||||  
QY 2401 GCCCATCTACACGCCCTTACTGGTAGCGCAAGAGCACTAAGGTGCGCGCTGCGTATGCA 2460  
DB |||||  
QY 2461 GCCCAAGGGTATAGGTGCTTGTCTCTGAA CCGCTCGGTGCGCGCA CCGTAGGTTTCGGG 2520  
DB |||||  
QY 2461 GCCCAAGGGTATAGGTGCTTGTCTCTGAA CCGCTCGGTGCGCGCA CCGTAGGTTTCGGG 2520  
DB |||||  
QY 2521 GCGTATATGCTAAGGCACATGGTATCGACCTAACATCAACATCAAGACCGGGTAAAGACCATC 2580  
DB |||||  
QY 2521 GCGTATATGCTAAGGCACATGGTATCGACCTAACATCAACATCAAGACCGGGTAAAGACCATC 2580  
DB |||||  
QY 2581 ACCACGGGTGCCCATCACGTACTCCACCTATGGCAAGTTTCTTGGCAGCGGTGTTGC 2640  
DB |||||  
QY 2581 ACCACGGGTGCCCATCACGTACTCCACCTATGGCAAGTTTCTTGGCAGCGGTGTTGC 2640  
DB |||||  
QY 2641 TCTGGGGCGCTATGATACATCAATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
DB |||||  
QY 2641 TCTGGGGCGCTATGATACATCAATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
DB |||||  
QY 2701 ATCTGGGCATCGGCACAGTCTCGACCCAGCGAGAGCGCTGAGCGGCACTCGTCGTG 2760  
DB |||||  
QY 2701 ATCTGGGCATCGGCACAGTCTCGACCCAGCGAGAGCGCTGAGCGGCACTCGTCGTG 2760  
DB |||||  
QY 2761 CTCGCCACCGCTACGCTCCGGGATCGGTCAACCGTGCCACATCAAAACATCGAGAGGTG 2820  
DB |||||  
QY 2761 CTCGCCACCGCTACGCTCCGGGATCGGTCAACCGTGCCACATCAAAACATCGAGAGGTG 2820  
DB |||||  
QY 2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
DB |||||  
QY 2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
DB |||||  
QY 2881 AAGGGGGGAGGCACTCATTTCTGCAATCCAGAGAAATGTGATGAGCTCGCGCG 2940  
DB |||||  
QY 2881 AAGGGGGGAGGCACTCATTTCTGCAATCCAGAGAAATGTGATGAGCTCGCGCG 2940  
DB |||||  
QY 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTTACCGGGGCTTGATGATCCGTC 3000  
DB |||||  
QY 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTTACCGGGGCTTGATGATCCGTC 3000  
DB |||||  
QY 3001 ATACCAACTAGCGGAGAGCTCATTTGTGTAGCAACCGAGCGCTCTAAATGACGGGTTTACC 3060  
DB |||||  
QY 3001 ATACCAACTAGCGGAGAGCTCATTTGTGTAGCAACCGAGCGCTCTAAATGACGGGTTTACC 3060  
DB |||||  
QY 3061 GCGGATTTGCACTCAGTGATCGACTGCAATACATGTGTCAACCGAGAGAGTCACTCAGC 3120  
DB |||||  
QY 3061 GCGGATTTGCACTCAGTGATCGACTGCAATACATGTGTCAACCGAGAGAGTCACTCAGC 3120  
DB |||||  
QY 3121 CTGACCCGACCTTCAACATTTGAGACAGACCGCTGCGACAAAGACGGGTGTCACTGTCG 3180  
DB |||||  
QY 3121 CTGACCCGACCTTCAACATTTGAGACAGACCGCTGCGACAAAGACGGGTGTCACTGTCG 3180  
DB |||||  
QY 3181 CAGCGGAGGACGAGACTGTTAGGGGAGAGATGGGCATTTACAGGTTTGTGACTCCAGGA 3240  
DB |||||  
QY 3181 CAGCGGAGGACGAGACTGTTAGGGGAGAGATGGGCATTTACAGGTTTGTGACTCCAGGA 3240  
DB |||||  
QY 3241 GAACGGCCCTCGGCATGTTCCGATTCCTCGGTTCTGTGCGAGTCTATGACCGGGCTGT 3300  
DB |||||  
QY 3241 GAACGGCCCTCGGCATGTTCCGATTCCTCGGTTCTGTGCGAGTCTATGACCGGGCTGT 3300  
DB |||||  
QY 3301 GCTTGTACGAGCTCAGCGCCCGCGAGACCTCAGTTAGGTTGCGGGCTTACTTAAACACA 3360  
DB |||||  
QY 3301 GCTTGTACGAGCTCAGCGCCCGCGAGACCTCAGTTAGGTTGCGGGCTTACTTAAACACA 3360  
DB |||||  
QY 3361 CCAGGTTGCCGCTGTGCCAGGACCATCTGAGTTCCTGGAGAGCGTCTTTACAGGCTC 3420  
DB |||||  
QY 3361 CCAGGTTGCCGCTGTGCCAGGACCATCTGAGTTCCTGGAGAGCGTCTTTACAGGCTC 3420  
DB |||||

QY 3421 ACCACATAGACGCCCTTCTTCTCCAGACTAAGCAGGAGAGAGCAACTTCCCTAC 3480  
DB |||||  
QY 3421 ACCACATAGACGCCCTTCTTCTCCAGACTAAGCAGGAGAGAGCAACTTCCCTAC 3480  
DB |||||  
QY 3481 CTGTTAGCATACACAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
DB |||||  
QY 3481 CTGTTAGCATACACAGGCTACGGTGTGCGCCAGGGCTCAGGCTCCACCTCCATCGTGGGAC 3540  
DB |||||  
QY 3541 CAATGTGGAGTGTCTCATACGGCTAAAGCTTAAGCTACGCTGCA CGGGCMAACGCCCTGTGTG 3600  
DB |||||  
QY 3541 CAATGTGGAGTGTCTCATACGGCTAAAGCTTAAGCTACGCTGCA CGGGCMAACGCCCTGTGTG 3600  
DB |||||  
QY 3601 TATAGGCTGGAGCCGTTCAAAACGAGTCTACCAACACACCCCAATACCAATACATC 3660  
DB |||||  
QY 3601 TATAGGCTGGAGCCGTTCAAAACGAGTCTACCAACACACCCCAATACCAATACATC 3660  
DB |||||  
QY 3661 ATGTCATGCTATGTCGGTGTGCTGAGAGTGTCTACAGAGCACTGGGTGCTGTAGGCGGA 3720  
DB |||||  
QY 3661 ATGTCATGCTATGTCGGTGTGCTGAGAGTGTCTACAGAGCACTGGGTGCTGTAGGCGGA 3720  
DB |||||  
QY 3721 GTCCTAGAGCTCTGGCGCGGTATGCTGACAAACAGGAGCTGTGCTATTTGGGAGG 3780  
DB |||||  
QY 3721 GTCCTAGAGCTCTGGCGCGGTATGCTGACAAACAGGAGCTGTGCTATTTGGGAGG 3780  
DB |||||  
QY 3781 ATCATCTTGTCCGAAAGCGGCCATCATTTCCGACAGGAGTCTTTACCGGAGTTC 3840  
DB |||||  
QY 3781 ATCATCTTGTCCGAAAGCGGCCATCATTTCCGACAGGAGTCTTTACCGGAGTTC 3840  
DB |||||  
QY 3841 GATGAGATGGAAGAGTGCCTCTACACTCCCTTTACATCGAACAGGGAATCAGCTGCC 3900  
DB |||||  
QY 3841 GATGAGATGGAAGAGTGCCTCTACACTCCCTTTACATCGAACAGGGAATCAGCTGCC 3900  
DB |||||  
QY 3901 GAACAAATTCAAACAGAGGCAATCGGGTGTGCAAAACAGCCACCAAGCAGGAGGCT 3960  
DB |||||  
QY 3901 GAACAAATTCAAACAGAGGCAATCGGGTGTGCAAAACAGCCACCAAGCAGGAGGCT 3960  
DB |||||  
QY 3961 GCTCTCCGCTGGTGGAAATCCAGTGCAGGACCTCGAAGCTTCTGGGGAGCATATG 4020  
DB |||||  
QY 3961 GCTCTCCGCTGGTGGAAATCCAGTGCAGGACCTCGAAGCTTCTGGGGAGCATATG 4020  
DB |||||  
QY 4021 TGGAAATTCATACGCGGATCAATATTTAGCAGGCTTGTCCACTCTGCTTGGCAACCCC 4080  
DB |||||  
QY 4021 TGGAAATTCATACGCGGATCAATATTTAGCAGGCTTGTCCACTCTGCTTGGCAACCCC 4080  
DB |||||  
QY 4081 GCGATAGCATCACTGATGGCATTCACAGCTCTATACAGCCGCTCAACACCAACAT 4140  
DB |||||  
QY 4081 GCGATAGCATCACTGATGGCATTCACAGCTCTATACAGCCGCTCAACACCAACAT 4140  
DB |||||  
QY 4141 ACCCTCTGTTTAACTCTGGGGGATGGTGGCGCCCAACTTGTCTCCAGCGCT 4200  
DB |||||  
QY 4141 ACCCTCTGTTTAACTCTGGGGGATGGTGGCGCCCAACTTGTCTCCAGCGCT 4200  
DB |||||  
QY 4201 GCTTCTCTTTCTGTTAGCGCGCATCGCTGGAGCGCTGTGTCAGCATAGGCTTTGGG 4260  
DB |||||  
QY 4201 GCTTCTCTTTCTGTTAGCGCGCATCGCTGGAGCGCTGTGTCAGCATAGGCTTTGGG 4260  
DB |||||  
QY 4261 AAGTGTCTGTGATATTTTGGCAGTTATGAGCAGGGTGGCAGCGGCTGTGTCGCC 4320  
DB |||||  
QY 4261 AAGTGTCTGTGATATTTTGGCAGTTATGAGCAGGGTGGCAGCGGCTGTGTCGCC 4320  
DB |||||  
QY 4321 TTTAAGGTCATAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTACTTCCCTGCT 4380  
DB |||||  
QY 4321 TTTAAGGTCATAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTACTTCCCTGCT 4380  
DB |||||  
QY 4381 ATCTCTCCCTCGCGCCCTAGTGTGTCGGGTGTGTCAGCAGCATACCTGCTGGCAG 4440  
DB |||||  
QY 4381 ATCTCTCCCTCGCGCCCTAGTGTGTCGGGTGTGTCAGCAGCATACCTGCTGGCAG 4440  
DB |||||  
QY 4441 GTGGGCCAGGAGAGGGGCTGTGCACTGTATGAA CCGGCTGTAGCGTTTGGCTTCGGG 4500  
DB |||||  
QY 4441 GTGGGCCAGGAGAGGGGCTGTGCACTGTATGAA CCGGCTGTAGCGTTTGGCTTCGGG 4500  
DB |||||  
QY 4501 GGTAAACACGCTCTCCCCCAGCATATGTGCTGAGAGCGGCTGCGACAGCTGTCACT 4560  
DB |||||

Db	4501	GTTAAACCAAGTCTCCCGGACGACTATGTCGCTGAGAGGAGCGCTGACGACGCTGTCACCT	4560	Db	5581	CCCATATGGGACGCGCCGGAATTAACACCTCCACCTGTTAGAGTCTCTGGAAGGACCCGGAC	5640
QY	4561	CAGATCCTCTTAGTCTTACCATCACTAGCTGCTGAAGAGGTTTCCAGTGGATCAAC	4620	QY	5641	TAGTTCCTTCCAGTGGTACACGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA	5700
Db	4561	CAGATCCTCTTAGTCTTACCATCACTAGCTGCTGAAGAGGTTTCCAGTGGATCAAC	4620	Db	5641	TAGTTCCTTCCAGTGGTACACGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA	5700
QY	4621	GAGGAGTCTCCACGCCATGCTCCGGCTGGTGGCTAAGAGATGTTTGGATTTGGATATGC	4680	QY	5701	CCTCCACGAGAGAAGGAGGAGGTTGCTGCTGAGAAATCTACCGTGTCTTCTGCTCTTGGCG	5760
Db	4621	GAGGAGTCTCCACGCCATGCTCCGGCTGGTGGCTAAGAGATGTTTGGATTTGGATATGC	4680	Db	5701	CCTCCACGAGAGAAGGAGGAGGTTGCTGCTGAGAAATCTACCGTGTCTTCTGCTCTTGGCG	5760
QY	4681	ACGGTGTGACTGATTTCAAGACTGCTCCAGTCCCAAGCTCTCGCGCGGATTTGCCGGA	4740	QY	5761	GAGCTCGCCACAAAAGACCTTCGCGAGCTCCGAATCTGCGCGCTCGACAGCGGCACGGCA	5820
Db	4681	ACGGTGTGACTGATTTCAAGACTGCTCCAGTCCCAAGCTCTCGCGCGGATTTGCCGGA	4740	Db	5761	GAGCTCGCCACAAAAGACCTTCGCGAGCTCCGAATCTGCGCGCTCGACAGCGGCACGGCA	5820
QY	4741	GTCCCTCTTCTCATGTCAATGTCAAGTCAAGTCAAGGAGTCTGGCGGGCGACGCGCATATG	4800	QY	5821	AGGCCCTCTCTGACACAGCCCTCCGACGAGGGGAGCGGGGATCCGACGTTGAGTCTGTAC	5880
Db	4741	GTCCCTCTTCTCATGTCAATGTCAAGTCAAGTCAAGGAGTCTGGCGGGCGACGCGCATATG	4800	Db	5821	AGGCCCTCTCTGACACAGCCCTCCGACGAGGGGAGCGGGGATCCGACGTTGAGTCTGTAC	5880
QY	4801	CAAAACCACTGCCATGTGGAGCACAGATCACCGGACATCTGAAACCGTTCATGAGG	4860	QY	5881	TCCTCCATGCCCCCTTGAGGGGAGCGGGGGATCCCGATCTCAGCGACGCGGTCTTGG	5940
Db	4801	CAAAACCACTGCCATGTGGAGCACAGATCACCGGACATCTGAAACCGTTCATGAGG	4860	Db	5881	TCCTCCATGCCCCCTTGAGGGGAGCGGGGGATCCCGATCTCAGCGACGCGGTCTTGG	5940
QY	4861	ATCGTGGGCTAGGACCTGTAGTAACACGTGCATGGAACATTCGCCATTTAACGGGTAC	4920	QY	5941	TCCTACCGTAAGCGAGGAGCTAGTGAAGACGTCGTCGCTGCTGCTGCTGCTACACATGG	6000
Db	4861	ATCGTGGGCTAGGACCTGTAGTAACACGTGCATGGAACATTCGCCATTTAACGGGTAC	4920	Db	5941	TCCTACCGTAAGCGAGGAGCTAGTGAAGACGTCGTCGCTGCTGCTGCTGCTACACATGG	6000
QY	4921	ACCACGGGCTTGCACGCGCTCCCGGGGCCAAATTAATCTAGGCGGTGTGGCGGGTG	4980	QY	6001	ACAGGCGCCTGATCAGCCATGCGCTGCGGAGGAAACCAAGTGCCTCATTAATGACATG	6060
Db	4921	ACCACGGGCTTGCACGCGCTCCCGGGGCCAAATTAATCTAGGCGGTGTGGCGGGTG	4980	Db	6001	ACAGGCGCCTGATCAGCCATGCGCTGCGGAGGAAACCAAGTGCCTCATTAATGACATG	6060
QY	4981	GCTGCTGAGAGTACGTGAGGTTACCGGGTGGGGATTTCCACATGCGGAGCGGATG	5040	QY	6061	AGCAACTCTTGTCTCGTCCGTCACCAACTTTGGTCTATGCTACAACTCTCGACGCGCAAGC	6120
Db	4981	GCTGCTGAGAGTACGTGAGGTTACCGGGTGGGGATTTCCACATGCGGAGCGGATG	5040	Db	6061	AGCAACTCTTGTCTCGTCCGTCACCAACTTTGGTCTATGCTACAACTCTCGACGCGCAAGC	6120
QY	5041	ACCACCTGACAAAGTAAAGTCCCGTGTCAAGTTCCCGGCCCGGAAATTTCTTCAAGAGTG	5100	QY	6121	CTCGGCGAGAAGAGTCACTTTTGACAGACTGCAAGTCTCTGCAAGCACTACCGGGAC	6180
Db	5041	ACCACCTGACAAAGTAAAGTCCCGTGTCAAGTTCCCGGCCCGGAAATTTCTTCAAGAGTG	5100	Db	6121	CTCGGCGAGAAGAGTCACTTTTGACAGACTGCAAGTCTCTGCAAGCACTACCGGGAC	6180
QY	5101	GATGGGTTGGTTGACAGGTACGTTCAAGTTCGCGGCGGCAACCCCTCTACGGGAGGTC	5160	QY	6181	GTGCTCAAGAGATGAGGAGGAGGCTCCACAGTTAAGCTTAACTTCTATCCGTTGGAG	6240
Db	5101	GATGGGTTGGTTGACAGGTACGTTCAAGTTCGCGGCGGCAACCCCTCTACGGGAGGTC	5160	Db	6181	GTGCTCAAGAGATGAGGAGGAGGCTCCACAGTTAAGCTTAACTTCTATCCGTTGGAG	6240
QY	5161	ACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGTCCCGGAGCCGGA	5220	QY	6241	GAGCCCTGTAAGCTGACGCGCCCACTTCGCGCGAGATCTTAAATTTGGCTATGGGCAAG	6300
Db	5161	ACATTCCTGCTGGGCTCAATCAATACCTGTTGGTTCACAGTCCCGGAGCCGGA	5220	Db	6241	GAGCCCTGTAAGCTGACGCGCCCACTTCGCGCGAGATCTTAAATTTGGCTATGGGCAAG	6300
QY	5221	CCGAGCTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACGGCGGAGCG	5280	QY	6301	GACGTCGGGAACTATCCAGCAAGGCGGTTAACCACTCCGCTCCGCTGTTGGAAGGACTTG	6360
Db	5221	CCGAGCTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACGGCGGAGCG	5280	Db	6301	GACGTCGGGAACTATCCAGCAAGGCGGTTAACCACTCCGCTCCGCTGTTGGAAGGACTTG	6360
QY	5281	GCTAAGCGTAGGCTGCGCGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG	5340	QY	6361	CTGGAAGACACTGAGACCAATTTGACACCACTCATGGCAAAATGAGGTTTCTGC	6420
Db	5281	GCTAAGCGTAGGCTGCGCGGGATCTCCCGCTCTTGGCCAGCTCATCAGCTAGCCAG	5340	Db	6361	CTGGAAGACACTGAGACCAATTTGACACCACTCATGGCAAAATGAGGTTTCTGC	6420
QY	5341	CTGCTGCGCTTCTTTGAGGCAATGCACTACCGCTACCTGCTACCTCCCGGAGCTGAC	5400	QY	6421	GTCCAAACAGAGAAAGGGGCGCGCAAGCCAGCTTCGCTTATCGTATTCGCAATTTGGGG	6480
Db	5341	CTGCTGCGCTTCTTTGAGGCAATGCACTACCGCTACCTGCTACCTCCCGGAGCTGAC	5400	Db	6421	GTCCAAACAGAGAAAGGGGCGCGCAAGCCAGCTTCGCTTATCGTATTCGCAATTTGGGG	6480
QY	5401	CTCATCGAGCCAACTCTGCTGGCGGAGGATGGGGGGAACATCACTCCCGCTGGAG	5460	QY	6481	GTTCTGCTGTGGAGAAAATGGCCCTTACGATGTGGTCTCCACCTCCCTCAGGCGCTG	6540
Db	5401	CTCATCGAGCCAACTCTGCTGGCGGAGGATGGGGGGAACATCACTCCCGCTGGAG	5460	Db	6481	GTTCTGCTGTGGAGAAAATGGCCCTTACGATGTGGTCTCCACCTCCCTCAGGCGCTG	6540
QY	5461	TCAGAAAATAGGTAGTAATTTTGGACTCTTTTCGAGCGCTCCAGCGGAGGATGAG	5520	QY	6541	ATGGGCTCTTCATACGGATTCNAATCTCTCTGGACAGGGGTGAGTCTCTGTTGAAT	6600
Db	5461	TCAGAAAATAGGTAGTAATTTTGGACTCTTTTCGAGCGCTCCAGCGGAGGATGAG	5520	Db	6541	ATGGGCTCTTCATACGGATTCNAATCTCTCTGGACAGGGGTGAGTCTCTGTTGAAT	6600
QY	5521	AGGGAAGTATCCGTTCCGGCGGAGATCCCTGCGAGGTCCAGGAAATTCCTCCAGGATG	5580	QY	6601	GCTCGAAAACGAGAAAATGCCCCCTATGGCTTCGCAATATGACACCCGCTGTTTGAUTCA	6660
Db	5521	AGGGAAGTATCCGTTCCGGCGGAGATCCCTGCGAGGTCCAGGAAATTCCTCCAGGATG	5580	Db	6601	GCTCGAAAACGAGAAAATGCCCCCTATGGCTTCGCAATATGACACCCGCTGTTTGAUTCA	6660
QY	5581	CCCATATGGGACGCGCCGGAATTAACACCTCCATCTGTTAGAGTCTCTGGAAGGACCCGGAC	5640	QY	6661	ACGGTCACTGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATGTTGAGTCTGGCC	6720
Db	5581	CCCATATGGGACGCGCCGGAATTAACACCTCCATCTGTTAGAGTCTCTGGAAGGACCCGGAC	5640	Db	6661	ACGGTCACTGAGAAATGACATCCGTTGAGGAGTCAATCTACCAATGTTGAGTCTGGCC	6720

```
QY 6721 CCCGAAGCCAGACAGGCGCATAGGCTCGCTACAGAGCGGCTTTACATCGGGGCCCCCTG 6780
Db 6721 CCCGAAGCCAGACAGGCGCATAGGCTCGCTACAGAGCGGCTTTACATCGGGGCCCCCTG 6780
QY 6781 ACTAATTCTAAAGGCGAGAACTGGGGCTATCGCGGTCGCCGCGGAGCGGTGTACTGACG 6840
Db 6781 ACTAATTCTAAAGGCGAGAACTGGGGCTATCGCGGTCGCCGCGGAGCGGTGTACTGACG 6840
QY 6841 ACCAGCTGCGGTAAATACCCCTCACATGTGTTACTGTGAAGCGCGTGTGCGAGCTGCG 6900
Db 6841 ACCAGCTGCGGTAAATACCCCTCACATGTGTTACTGTGAAGCGCGTGTGCGAGCTGCG 6900
QY 6901 AAGCTCCAGGACTCGACGATCGTGTATGCGGAGACGACCTGTGTTATCTGTGAAGC 6960
Db 6901 AAGCTCCAGGACTCGACGATCGTGTATGCGGAGACGACCTGTGTTATCTGTGAAGC 6960
QY 6961 GCGGGGACCCAAAGAGGACGAGCGGAGCTTACGGGCTTACGGAGGCTATGACTAGATAC 7020
Db 6961 GCGGGGACCCAAAGAGGAGGAGGAGCTTACGGGCTTACGGAGGCTATGACTAGATAC 7020
QY 7021 TCTGCCCCCTGGGAGCCGCCCAACAGAAATACGACTTGGAGTTGATTAACATCATGC 7080
Db 7021 TCTGCCCCCTGGGAGCCGCCCAACAGAAATACGACTTGGAGTTGATTAACATCATGC 7080
QY 7081 TCCTCCAAATGTGTCAGTCGCGCAGATGTCATCTGGCAAAAGGTTGACTATCTCACCGT 7140
Db 7081 TCCTCCAAATGTGTCAGTCGCGCAGATGTCATCTGGCAAAAGGTTGACTATCTCACCGT 7140
QY 7141 GACCCCAACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
Db 7141 GACCCCAACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
QY 7201 TCCTGGCTTAGGCAACATCATGATGTCGCCACCTTGTGGCAAGGATGATCCTGATG 7260
Db 7201 TCCTGGCTTAGGCAACATCATGATGTCGCCACCTTGTGGCAAGGATGATCCTGATG 7260
QY 7261 ACTCATTTCTTCTCCACCTCTAGCTCAGAAACAATTGAAAGCCCTAGATTGTCAG 7320
Db 7261 ACTCATTTCTTCTCCACCTCTAGCTCAGAAACAATTGAAAGCCCTAGATTGTCAG 7320
QY 7321 ATCTACGGGGCTGTACTCCATTGAGCACTTGACCTTACCTCAGATCAATCAACGACTC 7380
Db 7321 ATCTACGGGGCTGTACTCCATTGAGCACTTGACCTTACCTCAGATCAATCAACGACTC 7380
QY 7381 CATGGCTTAGCGATTTACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440
Db 7381 CATGGCTTAGCGATTTACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440
QY 7441 TCATGCGCTCAGGAACCTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAGT 7500
Db 7441 TCATGCGCTCAGGAACCTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAGT 7500
QY 7501 GTCCGCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTTC 7560
Db 7501 GTCCGCGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACTCTTTC 7560
QY 7561 AACTGGGAGTAAAGCAACAGCTCAAACTCACTCCAAATCCCGGCTGGTCCAGTTGGAT 7620
Db 7561 AACTGGGAGTAAAGCAACAGCTCAAACTCACTCCAAATCCCGGCTGGTCCAGTTGGAT 7620
QY 7621 TTATTCAGCTGGTTCGTTGCTGTATGAGCGGGGAGACATATATCACAGCTGTCTCGT 7680
Db 7621 TTATTCAGCTGGTTCGTTGCTGTATGAGCGGGGAGACATATATCACAGCTGTCTCGT 7680
QY 7681 GCCCGACCCCGCTGTTTCATGTGTGCTTACTCTACTTTCTGTAGGGGTAGGCACTCAT 7740
Db 7681 GCCCGACCCCGCTGTTTCATGTGTGCTTACTCTACTTTCTGTAGGGGTAGGCACTCAT 7740
QY 7741 CTACTCCCAACCGATGAACGGGAGCTTAAACACTTCCAGGCCAATAGGCCATCTCTGTTT 7800
Db 7741 CTACTCCCAACCGATGAACGGGAGCTTAAACACTTCCAGGCCAATAGGCCATCTCTGTTT 7800
```

```
QY 7801 TTTCCCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
Db 7801 TTTCCCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
QY 7861 TTTTTCCTCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7920
Db 7861 TTTTTCCTCTTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 7920
QY 7921 TAGCTGTGAAGGTCGCTGAGCCGCTTGACTGCGAGAGTCTGATATCGGCTCTCTGC 7980
Db 7921 TAGCTGTGAAGGTCGCTGAGCCGCTTGACTGCGAGAGTCTGATATCGGCTCTCTGC 7980
QY 7981 AGATCAAGT 7989
Db 7981 AGATCAAGT 7989

RESULT 2
US-10-005-469-1
; Sequence 1, Application US/10005469
; Publication No. US20020155133A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPI
; FILE REFERENCE: 0342/1H395US1
; CURRENT APPLICATION NUMBER: US/10/005,469
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV replicon 1377/NS3-3'UTR
US-10-005-469-1

Query Match 100.0%; Score 7987.4; DB 13; Length 7992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCCGATGGGGGCGACACTCCACATAGATCACTCCCTGTGAGAACTACTG 60
Db 1 GCCAGCCCCCGATGGGGGCGACACTCCACATAGATCACTCCCTGTGAGAACTACTG 60
QY 61 TCTTCACGCAAGACGCTAGCCATCGGCTTGTAGTGTGTCGAGCTCCAGAC 120
Db 61 TCTTCACGCAAGACGCTAGCCATCGGCTTGTAGTGTGTCGAGCTCCAGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTAGTACACCGGAATTCGCG 180
Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTAGTACACCGGAATTCGCG 180
QY 181 GACGACCGGTCCTTTCTTGGATCAACCCGCTCAATGCCTGAGATTGGCGCTGCCGCC 240
Db 181 GACGACCGGTCCTTTCTTGGATCAACCCGCTCAATGCCTGAGATTGGCGCTGCCGCC 240
QY 241 GCGAGACTGCTAGCCGATAGTGTGGTTCGGAAGGCCCTTGTGTTACTGCTCTGATAG 300
Db 241 GCGAGACTGCTAGCCGATAGTGTGGTTCGGAAGGCCCTTGTGTTACTGCTCTGATAG 300
QY 301 GTGCTTCGAGTGCCTCCCGGAGGTCTCGTAGCCGTCACCATGAGCACGAATCTTAAC 360
Db 301 GTGCTTCGAGTGCCTCCCGGAGGTCTCGTAGCCGTCACCATGAGCACGAATCTTAAC 360
QY 361 CTCAGAGAAAACCAAGGGCGGCCATGATTGAACAAGATTGGAATTCACGAGGTTCTC 420
Db 361 CTCAGAGAAAACCAAGGGCGGCCATGATTGAACAAGATTGGAATTCACGAGGTTCTC 420
QY 421 CGGCCGCTTGGGTGGAGAGGCTATTGGCTATGACTGGGCACACAGCAATCGGCTGCT 480
```

Db	421	CGSCCGCTTGGGTGGAGAGCTATTTCGGCTATGACTGGGCACACAGCAATCGGCTGCT	480	1501	AACCCCCACCTGGCGACAGGTGCTCTGCGGCCAAAAGCCACGCTGTATAAGATACACCT	1560
Qy	481	CTGATCGCGCGTGTTCGGCTGTACAGCGAGGGGCGCCGGTCTTTTGTCAAGACCG	540	1561	GCAAGGGCGGCACAAACCCAGTGCACACCTTGTGAGTTGATAGTTGTGGAAGAGTCAAA	1620
Db	481	CTGATCGCGCGTGTTCGGCTGTACAGCGAGGGGCGCCGGTCTTTTGTCAAGACCG	540	1561	GCAAGGGCGGCACAAACCCAGTGCACACCTTGTGAGTTGATAGTTGTGGAAGAGTCAAA	1620
Qy	541	ACCTGTCCGGTGCCTGAAATGAATCTCAGACACAGACGCGCGCTATCGTGCTGGCCA	600	1621	TGGCTCTCCTCAAGCGTATTCAACAAAGGGGCTGAAGATGCCACAGAGGTACCCCATTTGT	1680
Db	541	ACCTGTCCGGTGCCTGAAATGAATCTCAGACACAGACGCGCGCTATCGTGCTGGCCA	600	1621	TGGCTCTCCTCAAGCGTATTCAACAAAGGGGCTGAAGATGCCACAGAGGTACCCCATTTGT	1680
Qy	601	CGACGGGCGCTTCCTTGGCAGCTGTGTCTGACCTTGTCTACTGAAGCGGGAAGGCACTGGC	660	1681	ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTTACATGTGTTTGTAGTCGAGTTAAAA	1740
Db	601	CGACGGGCGCTTCCTTGGCAGCTGTGTCTGACCTTGTCTACTGAAGCGGGAAGGCACTGGC	660	1681	ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTTACATGTGTTTGTAGTCGAGTTAAAA	1740
Qy	661	TGCTATTGGCGAAGTGC CGGGGAGAGATCTCTGTCATCTCACTTGTCTCTGCGGAGA	720	1741	AAGCTCTAGCCCCCGAACCACCGGGACGTGGTTTTCTTTGAAAAACACGATAATACC	1800
Db	661	TGCTATTGGCGAAGTGC CGGGGAGAGATCTCTGTCATCTCACTTGTCTCTGCGGAGA	720	1741	AAGCTCTAGCCCCCGAACCACCGGGACGTGGTTTTCTTTGAAAAACACGATAATACC	1800
Qy	721	AAGTATCCATCATGGCTGATGCAATCGCGGCTGCATACGCTTGATTCGGCTACCTGCC	780	1801	ATGGCGCTATTACGGCTTACTCCCAACAGACCGGAGGCTACTTGGCTGCATCATCACT	1860
Db	721	AAGTATCCATCATGGCTGATGCAATCGCGGCTGCATACGCTTGATTCGGCTACCTGCC	780	1801	ATGGCGCTATTACGGCTTACTCCCAACAGACCGGAGGCTACTTGGCTGCATCATCACT	1860
Qy	781	CATTTCGACCAACAAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAAGCCGGTC	840	1861	AGCCTCACAGGCGGGACAGGAACAGGTTCGAGGGGGAGGTCCAAAGTGGTCTCCACCGCA	1920
Db	781	CATTTCGACCAACAAAGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGAAAGCCGGTC	840	1861	AGCCTCACAGGCGGGACAGGAACAGGTTCGAGGGGGAGGTCCAAAGTGGTCTCCACCGCA	1920
Qy	841	TTGTGATCAGATGATCTGGAGAAAGATCAGGGGCTCGCGCCAGCCGAACTGTTTCG	900	1921	ACAAATCTTTCTGGCGACCTGCGTCAATGGCGTGTGGACTGTCTATCATGTGTC	1980
Db	841	TTGTGATCAGATGATCTGGAGAAAGATCAGGGGCTCGCGCCAGCCGAACTGTTTCG	900	1921	ACAAATCTTTCTGGCGACCTGCGTCAATGGCGTGTGGACTGTCTATCATGTGTC	1980
Qy	901	CAAGGCTCAAGCGCGCATGCGCGAGAGATCTGCTGAGTATCGATGCGGATTCGCT	960	1981	GGCTCAAAGACCTTCCCGGCCCAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2040
Db	901	CAAGGCTCAAGCGCGCATGCGCGAGAGATCTGCTGAGTATCGATGCGGATTCGCT	960	1981	GGCTCAAAGACCTTCCCGGCCCAAGGGCCCAATCACCCAAATGTACACCAATGTGGAC	2040
Qy	961	GCTTTCGGAATCATGCTGGAAATGGCGCTTTTCTGGATTCATCGACTGGCGCGC	1020	2041	CAGGACCTCGTGGCGCAAGCGCCCCCGGGCGCGTTCCTTGAACCATGCACTGTC	2100
Db	961	GCTTTCGGAATCATGCTGGAAATGGCGCTTTTCTGGATTCATCGACTGGCGCGC	1020	2041	CAGGACCTCGTGGCGCAAGCGCCCCCGGGCGCGTTCCTTGAACCATGCACTGTC	2100
Qy	1021	TGGGTGTGGGACCCCTATCAGGATAGCGTGTGCTACCGGTGATTTGCTGAAGAC	1080	2101	GGCAGCTCGGACTTTTACTTGGTCAGAGGCAATGCGAGTGTCAATTCGGGTGCGCGCGG	2160
Db	1021	TGGGTGTGGGACCCCTATCAGGATAGCGTGTGCTACCGGTGATTTGCTGAAGAC	1080	2101	GGCAGCTCGGACTTTTACTTGGTCAGAGGCAATGCGAGTGTCAATTCGGGTGCGCGCGG	2160
Qy	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGTCCCGATTCCG	1140	2161	GGCGACAGAGGGGAGCTTACTCTCCCGAGCGCGCTCTCTTCTTGAAGGCTCTTCG	2220
Db	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGTCCCGATTCCG	1140	2161	GGCGACAGAGGGGAGCTTACTCTCCCGAGCGCGCTCTCTTCTTGAAGGCTCTTCG	2220
Qy	1141	AGCGCATCGCTTCTATCGCTTCTTGGAGTTCCTGAGTTTAAACAGACCAACACG	1200	2221	GGCGTCCACTGCTGTGCCCTCGGGGACAGTGTGGGCAATCTTTTGGGCTGCGGTGTC	2280
Db	1141	AGCGCATCGCTTCTATCGCTTCTTGGAGTTCCTGAGTTTAAACAGACCAACACG	1200	2221	GGCGTCCACTGCTGTGCCCTCGGGGACAGTGTGGGCAATCTTTTGGGCTGCGGTGTC	2280
Qy	1201	GTTCCTCTAGCGGATCAATTCGGCGCTCTCCCTCCCGCCCGCTTACGTTACTGCG	1260	2281	ACCGAGGGGTTCGAGAGGGGTGGACTTTGTACCGTTCGAGTCTATGGAACCACTATG	2340
Db	1201	GTTCCTCTAGCGGATCAATTCGGCGCTCTCCCTCCCGCCCGCTTACGTTACTGCG	1260	2281	ACCGAGGGGTTCGAGAGGGGTGGACTTTGTACCGTTCGAGTCTATGGAACCACTATG	2340
Qy	1261	CGAAGCGCTTGAATAAGCGCGTGTGCTTGTCTATATGTTTTCACCATATTG	1320	2341	CGGTCCCGGCTCTTCAACGACAACTCGTCCCTCCCGCGCTACCGCAGACATTCAGGTG	2400
Db	1261	CGAAGCGCTTGAATAAGCGCGTGTGCTTGTCTATATGTTTTCACCATATTG	1320	2341	CGGTCCCGGCTCTTCAACGACAACTCGTCCCTCCCGCGCTACCGCAGACATTCAGGTG	2400
Qy	1321	CGCTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTTCAGGACATTCCT	1380	2401	GCCCATCTACAGCCCTTACTGGTAGCGGCAAGAGCACTAAGGTGCGCGTGGTATGCA	2460
Db	1321	CGCTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTTCAGGACATTCCT	1380	2401	GCCCATCTACAGCCCTTACTGGTAGCGGCAAGAGCACTAAGGTGCGCGTGGTATGCA	2460
Qy	1381	AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGGTCTGTGAATGTGCTGAAGGAGCA	1440	2461	GCCCAAGGTATTAAGGTGCTTGTCTGAACCGCTCGTCCCGCCACCTTAGGTTTCGGG	2520
Db	1381	AGGGGTCTTTCCTCTCGCCAAAGGAATGCAAGGTCTGTGAATGTGCTGAAGGAGCA	1440	2461	GCCCAAGGTATTAAGGTGCTTGTCTGAACCGCTCGTCCCGCCACCTTAGGTTTCGGG	2520
Qy	1441	GTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCCCTTTTCAGGCGAGCGG	1500	2521	GGGTATATGCTTAAGGCACATGGTATCGACCCCTTAACATCAGAACCGGGGTAAAGCACATC	2580
Db	1441	GTTCCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCCCTTTTCAGGCGAGCGG	1500	2521	GGGTATATGCTTAAGGCACATGGTATCGACCCCTTAACATCAGAACCGGGGTAAAGCACATC	2580
Qy	1501	AACCCCCACCTGGCGACAGGTGCTCTGCGGCCAAAAGCCACGCTGTATAAGATACACCT	1560	2581	ACCAAGGGTGCCTGCGACAGGTGCTCTGCGGCCAAAAGCCACGCTGTATAAGATACACCT	2640



2641 TCTGGGGCGCCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
2641 TCTGGGGCGCCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
2701 ATCTGTGGGCATCGGCACAGCTCTTGGAACAAGCGAGAGCGGTGAGCGCGACTCGTGGT 2760  
2701 ATCTGTGGGCATCGGCACAGCTCTTGGAACAAGCGAGAGCGGTGAGCGCGACTCGTGGT 2760  
2761 CTGCGCCACCGGTACGCTTCGGGATCGGTCAACCGTGCACATCAAAACATCGAGAGGTG 2820  
2761 CTGCGCCACCGGTACGCTTCGGGATCGGTCAACCGTGCACATCAAAACATCGAGAGGTG 2820  
2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGCGAAGCAATCCCATCGAGACCATC 2880  
2821 GCTCTGTCCAGCACTGAGAAATCCCTTTTATGCGAAGCAATCCCATCGAGACCATC 2880  
2881 AAGGGGGGAGGACCTCATTTTCTGCCATTTCCAGAAGAAATGTGATGAGCTCGCCGCG 2940  
2881 AAGGGGGGAGGACCTCATTTTCTGCCATTTCCAGAAGAAATGTGATGAGCTCGCCGCG 2940  
2941 AAGCTGTCCGGCCTCGGACTCAATGTCTGTAGCATATTAACGGGGCCTTGATGTATCCGTC 3000  
2941 AAGCTGTCCGGCCTCGGACTCAATGTCTGTAGCATATTAACGGGGCCTTGATGTATCCGTC 3000  
3001 ATACCAACTAGCGAGAGCGTCAATGTCTGTAGCAAGCGGCTCTAATGACGGGCTTTACC 3060  
3001 ATACCAACTAGCGAGAGCGTCAATGTCTGTAGCAAGCGGCTCTAATGACGGGCTTTACC 3060  
3061 GGCATTTCCAGCTCAGTGTGCAATACATGTGTCAACCCAGACAGTGCAGCTTCAGC 3120  
3061 GGCATTTCCAGCTCAGTGTGCAATACATGTGTCAACCCAGACAGTGCAGCTTCAGC 3120  
3121 CTGGAACCGCATTTCAATTTAGACAGCAGCCGTGCGACAAAGACGGGTTCACGCTCG 3180  
3121 CTGGAACCGCATTTCAATTTAGACAGCAGCCGTGCGACAAAGACGGGTTCACGCTCG 3180  
3181 CAGCGGCGAGGACGAGTGTGTAGGGGAGGATGGGCATTTACAGGTTGTGACTCCAGGA 3240  
3181 CAGCGGCGAGGACGAGTGTGTAGGGGAGGATGGGCATTTACAGGTTGTGACTCCAGGA 3240  
3241 GAACGGCCCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTCTATGACGGGGTGT 3300  
3241 GAACGGCCCTCGGGCATGTTTCAATTCCTCGGTTCTGTGCGAGTGTCTATGACGGGGTGT 3300  
3301 GCTTGTACAGCTCAACCGCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
3301 GCTTGTACAGCTCAACCGCCGCGAGACCTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
3361 CCAGGTTGCCGCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGGGCTC 3420  
3361 CCAGGTTGCCGCTCTGCCAGGACCATCTGGAGTTCTGGGAGAGCGTCTTACAGGGCTC 3420  
3421 ACCACATAGACGCCATTTCTTCTCCAGACTAAGCAGGCGAGAGACAACTTCCCTTAC 3480  
3421 ACCACATAGACGCCATTTCTTCTCCAGACTAAGCAGGCGAGAGACAACTTCCCTTAC 3480  
3481 CTGTTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
3481 CTGTTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
3541 CAAATGTGGAGTGTCTCATACGCTAAAGCTACGCTAGCTGACGCGGCAACGCCCTCTGT 3600  
3541 CAAATGTGGAGTGTCTCATACGCTAAAGCTACGCTAGCTGACGCGGCAACGCCCTCTGT 3600  
3601 TATAGGCTGGAGCGCTTCAAAACGAGTTACTACCAACACACACCCCAATACCAATACATC 3660  
3601 TATAGGCTGGAGCGCTTCAAAACGAGTTACTACCAACACACACCCCAATACCAATACATC 3660  
3661 ATGSCATGCATGTGCGGTGACCTGGAGTGTCTCAGGACCACTGGGTGCTGGTAGGGGGA 3720  
3661 ATGSCATGCATGTGCGGTGACCTGGAGTGTCTCAGGACCACTGGGTGCTGGTAGGGGGA 3720

3721 GTCTAGCAGCTCTGGCGCGGTATTGCTGACAAAGCAGCGTGTCTATTGTGGCAGG 3780  
3721 GTCTAGCAGCTCTGGCGCGGTATTGCTGACAAAGCAGCGTGTCTATTGTGGCAGG 3780  
3781 ATCATCTTGTCCGAAAGCGGGCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
3781 ATCATCTTGTCCGAAAGCGGGCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
3841 GATGAGATGGAAGTGGCGCTCACCTCCCTTTACATCGAAAGGAAATCAGCTCGCC 3900  
3841 GATGAGATGGAAGTGGCGCTCACCTCCCTTTACATCGAAAGGAAATCAGCTCGCC 3900  
3901 GAACAAATTCACACAGAGGCAATCGGGTGTCTGCAACAGCACCAAGCAGCGAGGCT 3960  
3901 GAACAAATTCACACAGAGGCAATCGGGTGTCTGCAACAGCACCAAGCAGCGAGGCT 3960  
3961 GCTGCTCCCGTGTGGAATCCAGTGGCGGACCCCTCGAAGCCTTCTGGGCGAACAATG 4020  
3961 GCTGCTCCCGTGTGGAATCCAGTGGCGGACCCCTCGAAGCCTTCTGGGCGAACAATG 4020  
4021 TGGAAATTTCAACAGCGGATACAAATATTTAGAGGCTTTGCCACTCTGCTGTGGCAACCC 4080  
4021 TGGAAATTTCAACAGCGGATACAAATATTTAGAGGCTTTGCCACTCTGCTGTGGCAACCC 4080  
4081 GCGATAGCATCAGTGGGATTCACAGCCTCTATCACAGCCGCTTATCACAGCCGCTCACACCAACAT 4140  
4081 GCGATAGCATCAGTGGGATTCACAGCCTCTATCACAGCCGCTTATCACAGCCGCTCACACCAACAT 4140  
4141 ACCCTCTGTTTAAATCTCTGGGGGATGGTGGCGCGCCCAACTTGTCTCTCCAGCGCT 4200  
4141 ACCCTCTGTTTAAATCTCTGGGGGATGGTGGCGCGCCCAACTTGTCTCTCCAGCGCT 4200  
4201 GCTTCTCTTGTGTAGCGCGCGCATCGTGGAGCGGCTGTGGCAGCAGTAGGCGCTTGGG 4260  
4201 GCTTCTCTTGTGTAGCGCGCGCATCGTGGAGCGGCTGTGGCAGCAGTAGGCGCTTGGG 4260  
4261 AAGTGCTCTGGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGCGGCTCGTGCGC 4320  
4261 AAGTGCTCTGGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGCGGCTCGTGCGC 4320  
4321 TTTAAGGTCATAGCGCGGAGATGCCCTCCACCGAGCACTGTGTTAACTTCTCTGCT 4380  
4321 TTTAAGGTCATAGCGCGGAGATGCCCTCCACCGAGCACTGTGTTAACTTCTCTGCT 4380  
4381 ATCTCTTCCCTGTGCGCCCTAGTCTCGGGTGTGTGTGCGAGCGATATGCTGCTGGGAC 4440  
4381 ATCTCTTCCCTGTGCGCCCTAGTCTCGGGTGTGTGTGCGAGCGATATGCTGCTGGGAC 4440  
4441 GTGGGCGCGAGGAGGGGCTGTGCACTGAGTGAACCGGCTGTAGCGTTGCGTTCGCGG 4500  
4441 GTGGGCGCGAGGAGGGGCTGTGCACTGAGTGAACCGGCTGTAGCGTTGCGTTCGCGG 4500  
4501 GGTAAACCACTCTCTCCCGCAGCGCATATGTGCTGAGAGCGACGCTGAGCAGCTGTGCT 4560  
4501 GGTAAACCACTCTCTCCCGCAGCGCATATGTGCTGAGAGCGACGCTGAGCAGCTGTGCT 4560  
4561 CAGATCTCTCTAGTCTTTAGCATCACTCAGCTGTCTGAAGAGGCTTCAACAGTGGATCAAC 4620  
4561 CAGATCTCTCTAGTCTTTAGCATCACTCAGCTGTCTGAAGAGGCTTCAACAGTGGATCAAC 4620  
4621 GAGGACTGTCTCAGCGCATGCTCGGCTCGTGGCTAGAGATGTTTGGGATTTGATATGC 4680  
4621 GAGGACTGTCTCAGCGCATGCTCGGCTCGTGGCTAGAGATGTTTGGGATTTGATATGC 4680  
4681 ACGGTGTGACTGATTTTCAAGACTGCTGCTCAGTCCAAAGCTCTCTGCGCGCATTTGCCGGGA 4740  
4681 ACGGTGTGACTGATTTTCAAGACTGCTGCTCAGTCCAAAGCTCTCTGCGCGCATTTGCCGGGA 4740  
4741 GTCCCTTCTCTCATGTCAACGTGGGTCAAGAGGAGTCTGGGGGGCGAGCGCATCATG 4800  
4741 GTCCCTTCTCTCATGTCAACGTGGGTCAAGAGGAGTCTGGGGGGCGAGCGCATCATG 4800  
4801 CAACACCACTGCGCATGTGGAGCAGATCACCGGACATGTGMAAAACGGTTCCATGAGG 4860

Db 4801 CAAACCACTGCCCATGTGGAGCACAGATCACCGGACATGTGAAACAGGTTCATGAGG 4860  
 QY 4861 ATCGTGGGCTTAGGACCTGTAGTAACAGTGGCATGGAACATTCGCCATTAAACGGGTAC 4920  
 Db 4861 ATCGTGGGCTTAGGACCTGTAGTAACAGTGGCATGGAACATTCGCCATTAAACGGGTAC 4920  
 QY 4921 ACCACGGGCCCCGTGACGCCCTCCCGGCGCCAAATTTCTAGAGCGCTGTGCGGGGTG 4980  
 Db 4921 ACCACGGGCCCCGTGACGCCCTCCCGGCGCCAAATTTCTAGAGCGCTGTGCGGGGTG 4980  
 QY 4981 GCTGCTGAGGAGTACGTGAGGTTACGCGGGTGGGGATTTCCACTACGTGACGGGCATG 5040  
 Db 4981 GCTGCTGAGGAGTACGTGAGGTTACGCGGGTGGGGATTTCCACTACGTGAGCGGCAATG 5040  
 QY 5041 ACCACTGACAACTGAAGTCCCGTTCAGGTTCCCGGCCCGAATTTCTTCAGAGAATG 5100  
 Db 5041 ACCACTGACAACTGAAGTCCCGTTCAGGTTCCCGGCCCGAATTTCTTCAGAGAATG 5100  
 QY 5101 GATGGGCTGCGGTTCACAGTACGCTCCAGCGTCAAAACCCCTCTTACGGGAGGAGTTC 5160  
 Db 5101 GATGGGCTGCGGTTCACAGTACGCTCCAGCGTCAAAACCCCTCTTACGGGAGGAGTTC 5160  
 QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCCCGAA 5220  
 Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCCCGAA 5220  
 QY 5221 CCGGAGTAGAGTGTCACTTCATGCTCACCGACCCCTCCCACTTACCGGCGAGAGG 5280  
 Db 5221 CCGGAGTAGAGTGTCACTTCATGCTCACCGACCCCTCCCACTTACCGGCGAGAGG 5280  
 QY 5281 GCTAAGCTAGGCTGCGAGGGATCTCCGCCCTCTTGSCCAGCTCATCAGCTAGCCAG 5340  
 Db 5281 GCTAAGCTAGGCTGCGAGGGATCTCCGCCCTCTTGSCCAGCTCATCAGCTAGCCAG 5340  
 QY 5341 CTGCTGCGCTTCCCTTGAAGCAACATGCACTACCCGTCACTGACTCCCGGACGCTGAC 5400  
 Db 5341 CTGCTGCGCTTCCCTTGAAGCAACATGCACTACCCGTCACTGACTCCCGGACGCTGAC 5400  
 QY 5401 CTCATCGAGGCCAACCTCTGTGGCGCAGGAGATGGGGGGAACATCACCCCGTGGAG 5460  
 Db 5401 CTCATCGAGGCCAACCTCTGTGGCGCAGGAGATGGGGGGAACATCACCCCGTGGAG 5460  
 QY 5461 TCAGAAATAAGGTAGTAATTTTGGACTCTTCGAGCGCTCCAGCGGAGGAGTATGAG 5520  
 Db 5461 TCAGAAATAAGGTAGTAATTTTGGACTCTTCGAGCGCTCCAGCGGAGGAGTATGAG 5520  
 QY 5521 AGGGAAGTATCCCTTCGCGGAGATCCCTGCGGAGGTCCAGGAAATTCCTCGAGCGATG 5580  
 Db 5521 AGGGAAGTATCCCTTCGCGGAGATCCCTGCGGAGGTCCAGGAAATTCCTCGAGCGATG 5580  
 QY 5581 CCCATATGGCACGCCCGGATTACACCTCCACTGTAGAGTCCCTGGAGGACCCGGAC 5640  
 Db 5581 CCCATATGGCACGCCCGGATTACACCTCCACTGTAGAGTCCCTGGAGGACCCGGAC 5640  
 QY 5641 TAGCTCCCTCCAGTGTACACGGGTGTCCATTCGCGCTGCCAAGGCCCTCCGATACCA 5700  
 Db 5641 TAGCTCCCTCCAGTGTACACGGGTGTCCATTCGCGCTGCCAAGGCCCTCCGATACCA 5700  
 QY 5701 CTTCCACGAGGAGGAGGAGTGTCTGTCAAGATCTACCGTGTCTTCTGCTTGGCG 5760  
 Db 5701 CTTCCACGAGGAGGAGGAGTGTCTGTCAAGATCTACCGTGTCTTCTGCTTGGCG 5760  
 QY 5761 GAGCTGCCACAAAGACCTTCGAGCTCCGATCCGTCCGCGTCCAGACCGCACCGCA 5820  
 Db 5761 GAGCTGCCACAAAGACCTTCGAGCTCCGATCCGTCCGCGTCCAGACCGCACCGCA 5820  
 QY 5821 ACGGCTCTCTGACAGCCCTCCGACGACGGGATCCGAGCTTGAAGTCTGATC 5880  
 Db 5821 ACGGCTCTCTGACAGCCCTCCGACGACGGGATCCGAGCTTGAAGTCTGATC 5880  
 QY 5881 TCCTCATGCCCCCTTGAAGGGAGACCGGGGATCCCGATCTCAGCAGCGGCTTGG 5940

Db 5881 TCCTCATGCCCCCTTGAAGGGAGACCGGGGATCCCGATCTCAGCAGCGGTCTTGG 5940  
 QY 5941 TCTACCGTAAGCGAGAGAGCTAGTAGAGAGCTGCTGCTGCTCGATGCTCTACACATGG 6000  
 Db 5941 TCTACCGTAAGCGAGAGAGCTAGTAGAGAGCTGCTGCTGCTCGATGCTCTACACATGG 6000  
 QY 6001 ACAGGCGCTCTGATCAGCCATGCGCTGCGAGGAAACCAAGCTGCCATCAATGACATG 6060  
 Db 6001 ACAGGCGCTCTGATCAGCCATGCGCTGCGAGGAAACCAAGCTGCCATCAATGACATG 6060  
 QY 6061 AGCAACTCTTTGCTCCGTCAACCAACTTTGGTCTATGCTACAACTCTGCAACGCAAGC 6120  
 Db 6061 AGCAACTCTTTGCTCCGTCAACCAACTTTGGTCTATGCTACAACTCTGCAACGCAAGC 6120  
 QY 6121 CTGCGCAGAAAGAGTCACTTTGACAGACTGCAAGTCTGCAAGCAGCAGTACCGGGAC 6180  
 Db 6121 CTGCGCAGAAAGAGTCACTTTGACAGACTGCAAGTCTGCAAGCAGCAGTACCGGGAC 6180  
 QY 6181 GTGCTCAAGGAGATGAAGCGAAGGCGTCCACAGTTAAGGCTAAACCTTATCCGTGGAG 6240  
 Db 6181 GTGCTCAAGGAGATGAAGCGAAGGCGTCCACAGTTAAGGCTAAACCTTATCCGTGGAG 6240  
 QY 6241 GAAGCCTGTAAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGCTATGGGGCAAG 6300  
 Db 6241 GAAGCCTGTAAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGCTATGGGGCAAG 6300  
 QY 6301 GAGCTCCGAAACCTATCCAGCAGGCGGTTAACCATCGCTCCGCTGGAGAGGACTTG 6360  
 Db 6301 GAGCTCCGAAACCTATCCAGCAGGCGGTTAACCATCGCTCCGCTGGAGAGGACTTG 6360  
 QY 6361 CTGGAAGACACTGAGACCAATTTGACACACCATCATGGCAAAATGAGTCTTCTGC 6420  
 Db 6361 CTGGAAGACACTGAGACCAATTTGACACACCATCATGGCAAAATGAGTCTTCTGC 6420  
 QY 6421 GTCCAAACGAGAGAGGGGCGCAAGCAGCTCGCTTATCGTATTCACAGATTTGGG 6480  
 Db 6421 GTCCAAACGAGAGAGGGGCGCAAGCAGCTCGCTTATCGTATTCACAGATTTGGG 6480  
 QY 6481 GTTCGCTGTCGAGAGAAATGCCCTTTAGTGTGCTCCACCTCCCTCAGGCGGTG 6540  
 Db 6481 GTTCGCTGTCGAGAGAAATGCCCTTTAGTGTGCTCCACCTCCCTCAGGCGGTG 6540  
 QY 6541 ATGGGCTCTTCTATCCGATTTCCAACTCTCTGAGAGCGGTTCAGTCTCTGCTGAAT 6600  
 Db 6541 ATGGGCTCTTCTATCCGATTTCCAACTCTCTGAGAGCGGTTCAGTCTCTGCTGAAT 6600  
 QY 6601 GCTTGAAGAGCAAGAAATGCCCTATGGGCTTCGCAATGACACCCGCTGTTTGAATCA 6660  
 Db 6601 GCTTGAAGAGCAAGAAATGCCCTATGGGCTTCGCAATGACACCCGCTGTTTGAATCA 6660  
 QY 6661 ACAGTCACTGAGAAATGACATCCGCTGTGAGGAGTCAATCTACCAATGTTGACTTGGC 6720  
 Db 6661 ACAGTCACTGAGAAATGACATCCGCTGTGAGGAGTCAATCTACCAATGTTGACTTGGC 6720  
 QY 6721 CCCGAAGCAGACAGGCCATTAAGTCTCAGAGAGCGGCTTTACATCCGGGGCCCGCTG 6780  
 Db 6721 CCCGAAGCAGACAGGCCATTAAGTCTCAGAGAGCGGCTTTACATCCGGGGCCCGCTG 6780  
 QY 6781 ACTAATTTAAAGGCGAAGTTCGGCTATGCGCGTCCGCGCAGAGCGGTGTACTGACG 6840  
 Db 6781 ACTAATTTAAAGGCGAAGTTCGGCTATGCGCGTCCGCGCAGAGCGGTGTACTGACG 6840  
 QY 6841 ACCAGTCCGGTAATACCTCATGTTACTTGAAGCGCGCTGCGGCTGTGAGTGGC 6900  
 Db 6841 ACCAGTCCGGTAATACCTCATGTTACTTGAAGCGCGCTGCGGCTGTGAGTGGC 6900  
 QY 6901 AAGCTCCAGGACTGCAAGTGTCTGATGCGGAGACGACTTGTCTGTTCTGTGAAGC 6960  
 Db 6901 AAGCTCCAGGACTGCAAGTGTCTGATGCGGAGACGACTTGTCTGTTCTGTGAAGC 6960  
 QY 6961 GCGGGACCCAGAGAGACGAGGAGCCTTACCGGCTTACCGAGGCTATGACTAGATAC 7020  
 Db 6961 GCGGGACCCAGAGAGACGAGGAGCCTTACCGGCTTACCGAGGCTATGACTAGATAC 7020

QY 7021 TCTGCCCCCTGGGGCCCGCCAAACCCAGATACGACTTGGAGTTGATAACATATGC 7080  
 Db 7021 TCTGCCCCCTGGGGCCCGCCAAACCCAGATACGACTTGGAGTTGATAACATATGC 7080  
 QY 7081 TCTTCCCAATGTGTCAGTCCGCGCAGCATCTGCAAAAGGGGTACTATCTCACCGGT 7140  
 Db 7081 TCTTCCCAATGTGTCAGTCCGCGCAGCATCTGCAAAAGGGGTACTATCTCACCGGT 7140  
 QY 7141 GACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
 Db 7141 GACCCACACCCCTTGGCGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
 QY 7201 TCTTGGCTAGGCAACATCATGTATGCGGCCACCTTGTGGGCAAGGATCATCTGATG 7260  
 Db 7201 TCTTGGCTAGGCAACATCATGTATGCGGCCACCTTGTGGGCAAGGATCATCTGATG 7260  
 QY 7261 ACTCATTTCTTCTCATCTTCTAGCTCAGGAACAACCTTGAAGAGCCCTAGATTGTGAG 7320  
 Db 7261 ACTCATTTCTTCTCATCTTCTAGCTCAGGAACAACCTTGAAGAGCCCTAGATTGTGAG 7320  
 QY 7321 ATCTACGGGGCTGTACTTCATTTAGGCCAATTGACCTTACCTCAGATCAATCAACGACTC 7380  
 Db 7321 ATCTACGGGGCTGTACTTCATTTAGGCCAATTGACCTTACCTCAGATCAATCAACGACTC 7380  
 QY 7381 CATGSCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
 Db 7381 CATGSCCTTAGCGCATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT 7440  
 QY 7441 TCATGSCCTCAGGAACCTTGGGGTACCGCCCTTGGAGTCTCGAGACATCGGGCCAGAAAT 7500  
 Db 7441 TCATGSCCTCAGGAACCTTGGGGTACCGCCCTTGGAGTCTCGAGACATCGGGCCAGAAAT 7500  
 QY 7501 GTCCGGCTTAGGCTACTGTCAGGGGGGAGGGCTGCCAATTGFGCAAGTACCTCTTC 7560  
 Db 7501 GTCCGGCTTAGGCTACTGTCAGGGGGGAGGGCTGCCAATTGFGCAAGTACCTCTTC 7560  
 QY 7561 AACTGGGCAGTAAGGCAACACTCAACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 7620  
 Db 7561 AACTGGGCAGTAAGGCAACACTCAACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 7620  
 QY 7621 TTATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCACAGCTGTCCTGT 7680  
 Db 7621 TTATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCACAGCTGTCCTGT 7680  
 QY 7681 GCCCGACCCCGTGGTTCATGFGTGCCTACTCCTACTCTCTAGGGGTAGGCATCTAT 7740  
 Db 7681 GCCCGACCCCGTGGTTCATGFGTGCCTACTCCTACTCTCTAGGGGTAGGCATCTAT 7740  
 QY 7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800  
 Db 7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800  
 QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
 Db 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
 QY 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
 Db 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
 QY 7921 TAGCTGTGAAGGTCCGTGAGCCGCTTACCTGAGAGTGTCTGATCTGAGAGTGTCTGATCTGCTG 7980  
 Db 7921 TAGCTGTGAAGGTCCGTGAGCCGCTTACCTGAGAGTGTCTGATCTGAGAGTGTCTGATCTGCTG 7980  
 QY 7981 AGATCAAGT 7989  
 Db 7981 AGATCAAGT 7989

; Publication No. US20040005549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPI  
 ; FILE REFERENCE: 0342/1H395US3  
 ; CURRENT APPLICATION NUMBER: US/10/434,842  
 ; CURRENT FILING DATE: 2003-05-09  
 ; PRIOR APPLICATION NUMBER: US 10/233,307  
 ; PRIOR FILING DATE: 2002-08-28  
 ; PRIOR APPLICATION NUMBER: US 10/005,469  
 ; PRIOR FILING DATE: 2001-11-07  
 ; PRIOR APPLICATION NUMBER: US 60/245,866  
 ; PRIOR FILING DATE: 2000-11-07  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 1  
 ; LENGTH: 7992  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HCVlb-based chimeric replicon  
 ; US-10-434-842-1

Query Match 100.0%; Score 7987.4; DB 16; Length 7992;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GCCAGCCCCGATTGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAACTACTG 60  
 Db 1 GCCAGCCCCGATTGGGGCGACACTCCACATAGATCACTCCCTGTGAGGAACTACTG 60  
 QY 61 TCTTCAGCAGAAAAGCTCTAGCCATGGCGTTAGTATGAGTGTCTGAGAGCTCCAGAC 120  
 Db 61 TCTTCAGCAGAAAAGCTCTAGCCATGGCGTTAGTATGAGTGTCTGAGAGCTCCAGAC 120  
 QY 121 CCCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGAAATGCCAG 180  
 Db 121 CCCCCTCCCGGGAGAGCCATAGTGGTCTGCGGAACCGGTGAGTACACCGAAATGCCAG 180  
 QY 181 GACGACCGGGTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTGGGCGTGCCCC 240  
 Db 181 GACGACCGGGTCTTTCTTGGATCAACCCGCTCAATGCTGGAGATTGGGCGTGCCCC 240  
 QY 241 GCGAGACTGTAGCCGAGTAGTGTGGTCCGGAAGCCCTTGTGTTACTGCTGATAGG 300  
 Db 241 GCGAGACTGTAGCCGAGTAGTGTGGTCCGGAAGCCCTTGTGTTACTGCTGATAGG 300  
 QY 301 GTGCTTGGAGTGCCTCCGGAGGTCTGTAGACCGTGCACCATGAGCAGCAATCTTAAC 360  
 Db 301 GTGCTTGGAGTGCCTCCGGAGGTCTGTAGACCGTGCACCATGAGCAGCAATCTTAAC 360  
 QY 361 CTCAGGAAACCAACAAAGGGCGGCCATGATTGAACAAGATGGATTGCACGAGTTCTC 420  
 Db 361 CTCAGGAAACCAACAAAGGGCGGCCATGATTGAACAAGATGGATTGCACGAGTTCTC 420  
 QY 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCAACAACAGCAATCGGCTGC 480  
 Db 421 CGGCGCTTGGGTGAGAGGCTATTCCGCTATGACTGGGCAACAACAGCAATCGGCTGC 480  
 QY 481 CTGATCGCGCTGTCCGGCTGTCCGCTGTCAGCGAGGGCGCCGGTCTTTTCTCAAGACCG 540  
 Db 481 CTGATCGCGCTGTCCGGCTGTCCGCTGTCAGCGAGGGCGCCGGTCTTTTCTCAAGACCG 540  
 QY 541 ACCTGTCCGCTGCTGAATGAACCTGAGGACGAGGCGGCTATCGTGGCTGGCCA 600  
 Db 541 ACCTGTCCGCTGCTGAATGAACCTGAGGACGAGGCGGCTATCGTGGCTGGCCA 600  
 QY 601 CGACGGGCTTCTTGGCGAGCTGTGCTCGAGCTTGTCACTGAAGCGGGAAGGACTGGC 660  
 Db 601 CGACGGGCTTCTTGGCGAGCTGTGCTCGAGCTTGTCACTGAAGCGGGAAGGACTGGC 660  
 QY 661 TGCTATTGGGCGAGTCCCGGGCAGGATCTCCTGTCTCATCTCAGCTTGTCTCTCCGAGA 720

Db	661	TGCTATTGGGGGAAGTGGCGGGGAGGATCTCCTGTCACTCACTTGTCTCTCCGCGAGA	720
Qy	721	AGATATCCATCATGGGCTGATGCAATGGGGGGTGGATAGGCTTGATCCGGCTACCTGCC	780
Db	721	AGATATCCATCATGGGCTGATGCAATGGGGGGTGGATAGGCTTGATCCGGCTACCTGCC	780
Qy	781	CATTTCGACCAACCAAGCGAACAATCGCATCGAGCGAGCACGCTACTCGGATGGAAGCGGGTC	840
Db	781	CATTTCGACCAACCAAGCGAACAATCGCATCGAGCGAGCACGCTACTCGGATGGAAGCGGGTC	840
Qy	841	TTGTGATCAGGATGATGTGACGAAGAGCATCAGGGGCTCGCGCCAGCGGAATCTGTTCG	900
Db	841	TTGTGATCAGGATGATGTGACGAAGAGCATCAGGGGCTCGCGCCAGCGGAATCTGTTCG	900
Qy	901	CAAGGCTCAAGCGCGCATGCCGAGCGGAGAGATCTCGTCTGTGAACCATGGGCGATGCCCT	960
Db	901	CAAGGCTCAAGCGCGCATGCCGAGCGGAGAGATCTCGTCTGTGAACCATGGGCGATGCCCT	960
Qy	961	GCTTGGCGAATATCATGGTGGAAATGGCGGCTTTTCTGGATTCATCGACTGTGGCGGGC	1020
Db	961	GCTTGGCGAATATCATGGTGGAAATGGCGGCTTTTCTGGATTCATCGACTGTGGCGGGC	1020
Qy	1021	TGGGTGTGGCGACCGCTATCAGACATAGCTGTGGCTACCGTGATATGTCTGAAGAGC	1080
Db	1021	TGGGTGTGGCGACCGCTATCAGACATAGCTGTGGCTACCGTGATATGTCTGAAGAGC	1080
Qy	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCGGCTCCCGATTCGC	1140
Db	1081	TTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTACGGTATCGCGGCTCCCGATTCGC	1140
Qy	1141	AGCGCATCGGCTTCTATCGGCTTCTTGACGAGTTCCTCTGAGTTTAAACAGACCAACAAG	1200
Db	1141	AGCGCATCGGCTTCTATCGGCTTCTTGACGAGTTCCTCTGAGTTTAAACAGACCAACAAG	1200
Qy	1201	GTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCTTACGTTACTGGC	1260
Db	1201	GTTTCCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCTTACGTTACTGGC	1260
Qy	1261	CGAAGCGGCTTGAATAAGCGCGGTGCGTTTGTCTATATGTATTTTCCACCATATTG	1320
Db	1261	CGAAGCGGCTTGAATAAGCGCGGTGCGTTTGTCTATATGTATTTTCCACCATATTG	1320
Qy	1321	CCGTCTTTTGGCAATGTAGGGGCGGAAACCTGGCCCTGCTTCTTTGACGAGCATTCCT	1380
Db	1321	CCGTCTTTTGGCAATGTAGGGGCGGAAACCTGGCCCTGCTTCTTTGACGAGCATTCCT	1380
Qy	1381	AGGGGTCTTCCCTCTCGCAAGGAATCAAGGTCGTGTGAATGTCGTGAAGGAAGCA	1440
Db	1381	AGGGGTCTTCCCTCTCGCAAGGAATCAAGGTCGTGTGAATGTCGTGAAGGAAGCA	1440
Qy	1441	GTTTCCCTCTGAAGCTTCTTGAAGCAACAACCTCTGTAGCGACCTTTTCAGCGAGCGG	1500
Db	1441	GTTTCCCTCTGAAGCTTCTTGAAGCAACAACCTCTGTAGCGACCTTTTCAGCGAGCGG	1500
Qy	1501	AACCCGCCACTGGCGACAGGTGCTCTGGCGCAAAAGCCAGCTGTATGAAGATACACCT	1560
Db	1501	AACCCGCCACTGGCGACAGGTGCTCTGGCGCAAAAGCCAGCTGTATGAAGATACACCT	1560
Qy	1561	GCAAGCGGCAACAACCCAGTGCAGTGTGAGTTGGATAGTTGTGGAAGAGTCAAA	1620
Db	1561	GCAAGCGGCAACAACCCAGTGCAGTGTGAGTTGGATAGTTGTGGAAGAGTCAAA	1620
Qy	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT	1680
Db	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGAAGATGCCAGAGGTACCCCATTTGT	1680
Qy	1681	ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGGTTAAAA	1740
Db	1681	ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGGTTAAAA	1740
Qy	1741	AAGCTCTAGGCCCCCGAACAACGAGGAGCGTGGTTTCTTTTGAAGAAACACGATATACC	1800
Db	1741	AAGCTCTAGGCCCCCGAACAACGAGGAGCGTGGTTTCTTTTGAAGAAACACGATATACC	1800
Qy	1801	ATGGCGCCTATTACGGCCTACTCCCAACAGACGCGAGGCCCTACTTGGCTGCATCATCACT	1860
Db	1801	ATGGCGCCTATTACGGCCTACTCCCAACAGACGCGAGGCCCTACTTGGCTGCATCATCACT	1860
Qy	1861	AGCCTCACAGGCGCGGACAGGAACCAAGTGCAGAGGGGAGGTCCAAGTGTCTCCACCGCA	1920
Db	1861	AGCCTCACAGGCGCGGACAGGAACCAAGTGCAGAGGGGAGGTCCAAGTGTCTCCACCGCA	1920
Qy	1921	ACAAATCTTTCTGGGACCTGCGTCAATGGGCGTGTGTGGACTGTCTATCATGTGTGCC	1980
Db	1921	ACAAATCTTTCTGGGACCTGCGTCAATGGGCGTGTGTGGACTGTCTATCATGTGTGCC	1980
Qy	1981	GGCTCAAAGACCCCTTCGCGGCCAAAGGCCCAATCACCAAAATGTACCAATGTGGAC	2040
Db	1981	GGCTCAAAGACCCCTTCGCGGCCAAAGGCCCAATCACCAAAATGTACCAATGTGGAC	2040
Qy	2041	CAGGACCTCGTGGTGGCAAGCGCCCCGGGGCGGTTCTTTGACCAATGACACTGC	2100
Db	2041	CAGGACCTCGTGGTGGCAAGCGCCCCGGGGCGGTTCTTTGACCAATGACACTGC	2100
Qy	2101	GGCAGCTCGGACCTTTTACTTGTACAGAGGATGCCGATGTCAATTCGGTGCAGCGCGG	2160
Db	2101	GGCAGCTCGGACCTTTTACTTGTACAGAGGATGCCGATGTCAATTCGGTGCAGCGCGG	2160
Qy	2161	GGCGACAGCAGGGGGAGGCTTACTCTCCCGAGGCCGCTCTCTTCTTGAAGGCTCTTTCG	2220
Db	2161	GGCGACAGCAGGGGGAGGCTTACTCTCCCGAGGCCGCTCTCTTCTTGAAGGCTCTTTCG	2220
Qy	2221	GGCGGTCCTACTGTCTGCCCTCGGGGCGAGCTGTGGGCTCTTTGGGCTGGCGTGTGC	2280
Db	2221	GGCGGTCCTACTGTCTGCCCTCGGGGCGAGCTGTGGGCTCTTTGGGCTGGCGTGTGC	2280
Qy	2281	ACCGAGGGGTTGCGAAGCGGGTGGACTTTTGTACCGTTCGAGTCTATGGAACCACTATG	2340
Db	2281	ACCGAGGGGTTGCGAAGCGGGTGGACTTTTGTACCGTTCGAGTCTATGGAACCACTATG	2340
Qy	2341	CGGTCCCGGCTTCTACGAGCAACTCGTCCCTCCCGCGGTACCGAGACATTCACAGGTG	2400
Db	2341	CGGTCCCGGCTTCTACGAGCAACTCGTCCCTCCCGCGGTACCGAGACATTCACAGGTG	2400
Qy	2401	GCCCATCTACGCCCTTACTGGTAGCGGCAAGACATAAGGTGCGGCTCGGTATGCA	2460
Db	2401	GCCCATCTACGCCCTTACTGGTAGCGGCAAGACATAAGGTGCGGCTCGGTATGCA	2460
Qy	2461	GCCCAAGGTTAAGGTGCTTGTCTGAAACCGTTCGTCCCGCCACCTTAGGTTTCGGG	2520
Db	2461	GCCCAAGGTTAAGGTGCTTGTCTGAAACCGTTCGTCCCGCCACCTTAGGTTTCGGG	2520
Qy	2521	GGGTATATGTAAAGCAGATGGTATCGACCTTAAACATCAGAACCGGGTAAAGCACTC	2580
Db	2521	GGGTATATGTAAAGCAGATGGTATCGACCTTAAACATCAGAACCGGGTAAAGCACTC	2580
Qy	2581	ACCACGGGTGCCCCATCAGTACTCCACTATGGCAAGTTTCTTCGACGCGTGTTCG	2640
Db	2581	ACCACGGGTGCCCCATCAGTACTCCACTATGGCAAGTTTCTTCGACGCGTGTTCG	2640
Qy	2641	TCTTGGGGCGCTTACACATATAATATGTATGAGTGCACCTCACTGACTGCACACT	2700
Db	2641	TCTTGGGGCGCTTACACATATAATATGTATGAGTGCACCTCACTGACTGCACACT	2700
Qy	2701	ATCCTGGGATCGGACAGTCTTGGACCAAGCGGAGACGGCTGGAGCGGACTCGTCTG	2760
Db	2701	ATCCTGGGATCGGACAGTCTTGGACCAAGCGGAGACGGCTGGAGCGGACTCGTCTG	2760
Qy	2761	CTCGCAACCGCTACGCTCCGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG	2820
Db	2761	CTCGCAACCGCTACGCTCCGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG	2820
Qy	2821	GCTCTCTCAGACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880
Db	2821	GCTCTCTCAGACTCGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880

2881 AAGGGGGAGGACCTCATTTCTGCCATTCCAAAGAGAAATGTGATGAGCTCCCGG 2940  
2881 AAGGGGGAGGACCTCATTTCTGCCATTCCAAAGAGAAATGTGATGAGCTCCCGG 2940  
2941 AAGCTGTCGGGCTCGGACTCAATCTGTAGCATATTAACGGGGCTTGTATGTATCCGTC 3000  
2941 AAGCTGTCGGGCTCGGACTCAATCTGTAGCATATTAACGGGGCTTGTATGTATCCGTC 3000  
3001 ATACCAACTAGCGGAGAGCTGATCTGTGTAGCAACGGAGCGCTCTAATGACGGCTTTACC 3060  
3001 ATACCAACTAGCGGAGAGCTGATCTGTGTAGCAACGGAGCGCTCTAATGACGGCTTTACC 3060  
3061 GCGATTTTCGACTCAGTGTGATCGACTGCAATCATGTGTCAACGACAGCTGACTTCAGC 3120  
3061 GCGATTTTCGACTCAGTGTGATCGACTGCAATCATGTGTCAACGACAGCTGACTTCAGC 3120  
3121 CTGGACCCGACTTCAACATTGAGACGACGACCGTGCACAAAGACGCGGTGTACGCTCG 3180  
3121 CTGGACCCGACTTCAACATTGAGACGACGACCGTGCACAAAGACGCGGTGTACGCTCG 3180  
3181 CAGCGCGGAGGACGACTGTTAGGGGAGGATGGGCAATTACAGTTTGTGACTCCAGGA 3240  
3181 CAGCGCGGAGGACGACTGTTAGGGGAGGATGGGCAATTACAGTTTGTGACTCCAGGA 3240  
3241 GAAACGGCCCTCGGGCATGTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGCTGT 3300  
3241 GAAACGGCCCTCGGGCATGTTCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGCTGT 3300  
3301 GCTTGGTACGAGCTCACGCGCGCCGAGACCTCAGTTAGTTGCGGGCTTACTTAACACA 3360  
3301 GCTTGGTACGAGCTCACGCGCGCCGAGACCTCAGTTAGTTGCGGGCTTACTTAACACA 3360  
3361 CCAGGTTGCCGCTGTCGACGAGACCATCTGGAGTTCTGGGAGAGGCTTTACAGGCTC 3420  
3361 CCAGGTTGCCGCTGTCGACGAGACCATCTGGAGTTCTGGGAGAGGCTTTACAGGCTC 3420  
3421 ACCACATPAGACGCCATTTCTGTGCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC 3480  
3421 ACCACATPAGACGCCATTTCTGTGCCAGACTAAGCAGGAGGAGACAACTTCCCTTAC 3480  
3481 CTGTTAGCATACAGGCTACGGTGTGCCAGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
3481 CTGTTAGCATACAGGCTACGGTGTGCCAGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
3541 CAATGTGGAGTGTCTCATAGGCTAAGCTTACGCTGACGGGCAACGCCCTCGT 3600  
3541 CAATGTGGAGTGTCTCATAGGCTAAGGCTTACGCTGACGGGCAACGCCCTCGT 3600  
3601 TATAGGCTGGGAGCGGTTCAAAAAGAGTTTACTACACACACCCCATAAACCAATACATC 3660  
3601 TATAGGCTGGGAGCGGTTCAAAAAGAGTTTACTACACACACCCCATAAACCAATACATC 3660  
3661 ATGGCATGCAATGTGCGGTGACCTGAGTGTGCTGACGACACCTGAGTGTGTTAGGCGGA 3720  
3661 ATGGCATGCAATGTGCGGTGACCTGAGTGTGCTGACGACACCTGAGTGTGTTAGGCGGA 3720  
3721 GTCCTAGCACTGTGGCGGCTATGCTGTGACACAGCAGCAGTGTGATGTGGGAGG 3780  
3721 GTCCTAGCACTGTGGCGGCTATGCTGTGACACAGCAGCAGTGTGATGTGGGAGG 3780  
3781 ATCATCTTGTCCGAAAGCGGCCATCATTTCCGACAGGGAAGTCTTTTACCGGGAGTTC 3840  
3781 ATCATCTTGTCCGAAAGCGGCCATCATTTCCGACAGGGAAGTCTTTTACCGGGAGTTC 3840  
3841 GATGAGATGGAAGTGGCGCTCACCTCCCTTACATCGAACAGGGAATCAGCTCGCC 3900  
3841 GATGAGATGGAAGTGGCGCTCACCTCCCTTACATCGAACAGGGAATCAGCTCGCC 3900  
3901 GAACAATTCACACAGAGGCAATCGGTTGTGTGAAACAGGACCAACAGCAGGAGGCT 3960  
3901 GAACAATTCACACAGAGGCAATCGGTTGTGTGAAACAGGACCAACAGCAGGAGGCT 3960  
3961 GCTGCTCCCGTGGTGGAAATCCAAGTGGGCGGCCCTCGAAGCCTTCTGGGCGAAGCATATG 4020

3961 GCTGCTCCCGTGGTGGAAATCCAAGTGGGCGACCCCTCGAAGCCCTTCTGGGCGAAGCATATG 4020  
4021 TGGAAATTTTCATCAGCGGATACAATATTTAGCAGGCTTTGTCCACTCTGCTGGCAACCC 4080  
4021 TGGAAATTTTCATCAGCGGATACAATATTTAGCAGGCTTTGTCCACTCTGCTGGCAACCC 4080  
4081 GCGATAGCATCACTGATGGCAATTCACGCTTCTATCACAGCCCGCTCACCACCAACAT 4140  
4081 GCGATAGCATCACTGATGGCAATTCACGCTTCTATCACAGCCCGCTCACCACCAACAT 4140  
4141 ACCCTCTGTTTAAATCCTCGGGGATGGGTGGCGCCCAACTTGTCTCCACAGCGCT 4200  
4141 ACCCTCTGTTTAAATCCTCGGGGATGGGTGGCGCCCAACTTGTCTCCACAGCGCT 4200  
4201 GCTTCTGCTTTCGTAGGCGCGGCATCGCTGAGCGGCTGTTGGCAGCATAGGCTTTGG 4260  
4201 GCTTCTGCTTTCGTAGGCGCGGCATCGCTGAGCGGCTGTTGGCAGCATAGGCTTTGG 4260  
4261 AAGTGTCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGCTCGTGCC 4320  
4261 AAGTGTCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGCTCGTGCC 4320  
4321 TTTAAGGTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 4380  
4321 TTTAAGGTCATGAGCGGAGATGCCCTCCACCGAGGACCTGGTTAACTACTCCCTGCT 4380  
4381 ATCTCTCTCCCTGGCGCCCTAGTCTCGGGGTCTGTGTCGAGCGCATATCTGCTCGGCAC 4440  
4381 ATCTCTCTCCCTGGCGCCCTAGTCTCGGGGTCTGTGTCGAGCGCATATCTGCTCGGCAC 4440  
4441 GTGGCCACAGGGAGGGGCTGTGAGTGTGATGAAACCGGCTGATAGCCTTGCCTTCGCG 4500  
4441 GTGGCCACAGGGAGGGGCTGTGAGTGTGATGAAACCGGCTGATAGCCTTGCCTTCGCG 4500  
4501 GGTAAACACAGCTCTCCCGACGACTATGTGCTCAGAGCGAGCTGACAGCGTGTCACT 4560  
4501 GGTAAACACAGCTCTCCCGACGACTATGTGCTCAGAGCGAGCTGACAGCGTGTCACT 4560  
4561 CAGATCTCTCTAGTCTTTACCATCACTCAGCTGTCTGAAGAGCTTTCACAGTGGATCAAC 4620  
4561 CAGATCTCTCTAGTCTTTACCATCACTCAGCTGTCTGAAGAGCTTTCACAGTGGATCAAC 4620  
4621 GAGACTGTCTCCACGCAATGCTCGGCTCGTGGCTTAAGAGATGTTTGGGATTTGATATGC 4680  
4621 GAGACTGTCTCCACGCAATGCTCGGCTCGTGGCTTAAGAGATGTTTGGGATTTGATATGC 4680  
4681 ACGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 4740  
4681 ACGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 4740  
4741 GTCCTCTTCTCTCATGTCAACGTTGAGTCAAGGAGTCTGCGCGGGCGAGCGCATCATG 4800  
4741 GTCCTCTTCTCTCATGTCAACGTTGAGTCAAGGAGTCTGCGCGGGCGAGCGCATCATG 4800  
4801 CAACACACCTGCTCCATGTGAGGACAGATCAACGCAATGTGAAAAACGTTTCCATGAGG 4860  
4801 CAACACACCTGCTCCATGTGAGGACAGATCAACGCAATGTGAAAAACGTTTCCATGAGG 4860  
4861 ATCGTGGGCGCTAGGACCTGTAGTAAACGCTGGCATGGAACATTTCCCATTAACGCGTAC 4920  
4861 ATCGTGGGCGCTAGGACCTGTAGTAAACGCTGGCATGGAACATTTCCCATTAACGCGTAC 4920  
4921 ACCACGGGCGCTTGCACGCGCTCCCGGGCGCAAAATTTATTTAGGCGGCTGTGGCGGGT 4980  
4921 ACCACGGGCGCTTGCACGCGCTCCCGGGCGCAAAATTTATTTAGGCGGCTGTGGCGGGT 4980  
4981 GCTGCTGAGGAGTACGTGGAGGTTAACCGGCTGGGGATTTCCACTACGTGACGGGATG 5040  
4981 GCTGCTGAGGAGTACGTGGAGGTTAACCGGCTGGGGATTTCCACTACGTGACGGGATG 5040  
5041 ACCACTGACAACTGTAAGTCCCGTGTACGTTCCGGCCCGCCGAAATTTCTTACAGAGTG 5100





Qy	7261	ACTCATTTCTTCCATCCTTTAGCTCAGAACAACTTGAAAGAAGCCCTAGATTGTTCAG	7320
Db	7261	ACTCATTTCTTCTCCATCCTTTAGCTCAGGAAACAATCTGAAAGAAGCCCTAGATTGTTCAG	7320
Qy	7321	ATCTACGGGGCGCTGTACTCATTTAGCACACTTCACCTACTCTCAGATCATTTCAACGACTC	7380
Db	7321	ATCTACGGGGCGCTGTACTCATTTGAGCCACTTGACCTTACCTCAGATCATTTCAACGACTC	7380
Qy	7381	CATGGCCCTTAGCGCATTTTTCACCTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT	7440
Db	7381	CATGGCCCTTAGCGCATTTTTCACCTCCATAGTTACTCTCCAGGTGAGATCAATAGGGTGGCT	7440
Qy	7441	TCATGCCCTCAGGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAGT	7500
Db	7441	TCATGCCCTCAGGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAGT	7500
Qy	7501	GTCGGCCTAGGCTACTGTCCACGGGGGGAGGGCTGCCACTGTGTGCAAGTACTCTTTC	7560
Db	7501	GTCGGCCTAGGCTACTGTCCACGGGGGGAGGGCTGCCACTGTGTGCAAGTACTCTTTC	7560
Qy	7561	AACTGGGCAGTAAGGACCAAGCTCAAACCTCACTCCAATCCCGGCTGGCTCCCAAGTTGGAT	7620
Db	7561	AACTGGGCAGTAAGGACCAAGCTCAAACCTCACTCCAATCCCGGCTGGCTCCCAAGTTGGAT	7620
Qy	7621	TTATCCAGCTGGTTGCTGTGTGTGTTA CAGCGGGGGAGACATATACACGCTGTCTCGT	7680
Db	7621	TTATCCAGCTGGTTGCTGTGTGTGTTA CAGCGGGGGAGACATATATACAGGCTGTCTCGT	7680
Qy	7681	GCCCGACCCCGCTGGTTTCATGTGGTGCTACTCTACTTTCTGTAGGGGTAGGCATCTAT	7740
Db	7681	GCCCGACCCCGCTGGTTTCATGTGGTGCTACTCTACTTTCTGTAGGGGTAGGCATCTAT	7740
Qy	7741	CTACTCCCCAACCGATGAAAGGGGACCTAAACACTCCAGGCCAATAGCCCATCTCTGTTTT	7800
Db	7741	CTACTCCCCAACCGATGAAAGGGGACCTAAACACTCCAGGCCAATAGCCCATCTCTGTTTT	7800
Qy	7801	TTTCCCTTT	7860
Db	7801	TTTCCCTTT	7860
Qy	7861	TTTTTCTCTCTTTTTTTTCCCTTTTCTTTCCCTTTGGTGGCTCCATCTTAGCCCTAGTCACGGC	7920
Db	7861	TTTTTCTCTCTTTTTTTTCCCTTTTCTTTCCCTTTGGTGGCTCCATCTTAGCCCTAGTCACGGC	7920
Qy	7921	TAGCTGTGAAGGTCCTGTAGCGCCTTTGACTGTGACAGAGTGTCTGATACTGSCCTCTCTCGC	7980
Db	7921	TAGCTGTGAAGGTCCTGTAGCGCCTTTGACTGTGACAGAGTGTCTGATACTGSCCTCTCTCGC	7980
Qy	7981	AGATCAAGT	7989
Db	7981	AGATCAAGT	7989

RESULT 4	
US-10-125-940-1	
; Sequence 1, Application US/10125940	
; Publication No. US20030045568A1	
; GENERAL INFORMATION:	
; APPLICANT: Altamura, Sergio	
; APPLICANT: Kock, Uwe	
; TITLE OF INVENTION: Treating Hepatitis C Viral Infections	
; TITLE OF INVENTION: with Thiosemicarbazone Compounds	
; FILE REFERENCE: IT0011Y	
; CURRENT APPLICATION NUMBER: US/10/125,940	
; CURRENT FILING DATE: 2002-05-07	
; PRIOR APPLICATION NUMBER: 60/285,104	
; PRIOR FILING DATE: 2001-04-20	
; NUMBER OF SEQ ID NOS: 1	
; SOFTWARE: FastSEQ for Windows Version 4.0	
; SEQ ID NO 1	
; LENGTH: 10690	
; TYPE: DNA	
; ORGANISM: Artificial Sequence	

  

QY	661	TGCTATTGGCGGAAGTGC	GCGGGCAGGATCCCTGTCA	TTCACCTTGCTCTGCCGAGA	720
Dd	661	TGCTATTGGCGGAAGTGC	GCGGGCAGGATCTCTGTCA	TTCACCTTGCTCTGCCGAGA	720
QY	721	AAGTATCCATCATGCTG	CATCGAATCGCGGGCTGC	ATAGCTTTGATCCGGCTAC	CTGCC
Dd	721	AAGTATCCATCATGCTG	CATCGAATCGCGGGCTGC	ATAGCTTTGATCCGGCTAC	CTGCC
QY	781	CATTTCGACCACCAAGC	GAACAATCGCATCGAGCG	CAGCACGTA	TCGGATCGAAAGCCGGTC
Dd	781	CATTTCGACCACCAAGC	GAACAATCGCATCGAGCG	CAGCACGTA	TCGGATCGAAAGCCGGTC
QY	841	TTGTGATCAGATGATCT	GACGAAGAGCATCAGGGG	CTCGGCCACGCCGAAC	TGTCG
Dd	841	TTGTGATCAGATGATCT	GACGAAGAGCATCAGGGG	CTCGGCCACGCCGAAC	TGTCG
QY	901	CCAGGCTCAAGCGCGCAT	CGCCGACCGCGAGGAT	CTCGTGGTAGCCATGGG	ATGCCT
Dd	901	CCAGGCTCAAGCGCGCAT	CGCCGACCGCGAGGAT	CTCGTGGTAGCCATGGG	ATGCCT
QY	961	GCCTTCCGAATATCATG	TGTGAAAATGCCCGCT	TTTTCTGGATT	CATCGACTGTGGCCGGC
					1020

Db	961	GCTTGCAGAAATATATGATGATGAAATAGCGCGCTTTCTTCTGATTCATCGACTGTGGCCGCG	1020
QY	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCGGTGGCTACCGGTGATATGCTGAAGAGC	1080
Db	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGCGGTGGCTACCGGTGATATGCTGAAGAGC	1080
QY	1081	TTGGCGGCAATGGGCTGACCGCTTCTCGTGGCTTTACGGTATCGCGCTCCCGATTGCG	1140
Db	1081	TTGGCGGCAATGGGCTGACCGCTTCTCGTGGCTTTACGGTATCGCGCTCCCGATTGCG	1140
QY	1141	AGCGCATCGCCCTTCTATCGCCCTTTCTGACGAGTTCTTCTGAGTTTAAACAGACCAACAAG	1200
Db	1141	AGCGCATCGCCCTTCTATCGCCCTTTCTGACGAGTTCTTCTGAGTTTAAACAGACCAACAAG	1200
QY	1201	GTTTCCCTCTAGGGGATCAATTTCCGCCCTCTCCCTCCGCCCTCCCGCTAACGTTACTGGC	1260
Db	1201	GTTTCCCTCTAGGGGATCAATTTCCGCCCTCTCCCTCCGCCCTCCCGCTAACGTTACTGGC	1260
QY	1261	CGAAGCGGCTTGAATTAAGCGGCTGTGGTGTGCTATATGTTATTTTCCACCATATTG	1320
Db	1261	CGAAGCGGCTTGAATTAAGCGGCTGTGGTGTGCTATATGTTATTTTCCACCATATTG	1320
QY	1321	CGCTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCT	1380
Db	1321	CGCTCTTTTGGCAATGTGAGGGCCCGGAAACCTGGCCCTGTCTTCTTGACGAGCATTCCT	1380
QY	1381	AGGGTCTTTCCCTCTCCGCAAGGAATGCAAGGCTGTGTTGATGTCTGTAAGGAAGCA	1440
Db	1381	AGGGTCTTTCCCTCTCCGCAAGGAATGCAAGGCTGTGTTGATGTCTGTAAGGAAGCA	1440
QY	1441	GTTCTCTGAGAGCTTTCTTGAAGACAAACAAGCTGTGAGCGACCTTTTGACGAGCGG	1500
Db	1441	GTTCTCTGAGAGCTTTCTTGAAGACAAACAAGCTGTGAGCGACCTTTTGACGAGCGG	1500
QY	1501	AACCCGCCCTCTGGGACAGGTGCTCTGGCGGCAAAAGCCAGTGTATAAGATACACCT	1560
Db	1501	AACCCGCCCTCTGGGACAGGTGCTCTGGCGGCAAAAGCCAGTGTATAAGATACACCT	1560
QY	1561	GCAAGGGGCAACACCCAGTGCACGTTGTGAGTTGATAGTTGTGGAAGAGTCAAA	1620
Db	1561	GCAAGGGGCAACACCCAGTGCACGTTGTGAGTTGATAGTTGTGGAAGAGTCAAA	1620
QY	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGGAAGGATGCCAGAGGTACCCCATTTG	1680
Db	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTGGAAGGATGCCAGAGGTACCCCATTTG	1680
QY	1681	ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTTACATGTGTTTGTAGTCGAGGTTAAAA	1740
Db	1681	ATGGGATCTGATCTGGGGCTCGGTGCACATGCTTTTACATGTGTTTGTAGTCGAGGTTAAAA	1740
QY	1741	AAGCTCTAGGCCCCCGAACACCGGGAGCTGGTGTTCCTTTGAAAAACACGATATACC	1800
Db	1741	AAGCTCTAGGCCCCCGAACACCGGGAGCTGGTGTTCCTTTGAAAAACACGATATACC	1800
QY	1801	ATGGGCTTATACGGCTTACTCCAAAGAGCGGAGGCTACTTGGCTGCATCATCACT	1860
Db	1801	ATGGGCTTATACGGCTTACTCCAAAGAGCGGAGGCTACTTGGCTGCATCATCACT	1860
QY	1861	AGCTCTACAGCGGGGACAGGAACAGGTTCGAGGGGAGGTCCAAAGTGTCTCCACCGCA	1920
Db	1861	AGCTCTACAGCGGGGACAGGAACAGGTTCGAGGGGAGGTCCAAAGTGTCTCCACCGCA	1920
QY	1921	ACAAATCTTTCTGGGACCTCGCTCAATGGGCTGTGTTGACTGTCTATCATGTGGC	1980
Db	1921	ACAAATCTTTCTGGGACCTCGCTCAATGGGCTGTGTTGACTGTCTATCATGTGGC	1980
QY	1981	GGCTCAAGAGCTTCCCGGCCCCAAAGGCCCCAATCACCCAAATGTACACCAATGTGGAC	2040
Db	1981	GGCTCAAGAGCTTCCCGGCCCCAAAGGCCCCAATCACCCAAATGTACACCAATGTGGAC	2040
QY	2041	CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGCTTCTTGTGACCAATGACCTGC	2100
Db	2041	CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGCTTCTTGTGACCAATGACCTGC	2100
QY	2101	GGCAGCTCGGACCTTTACTTGTCTCAGGAGCATGCGATGTCTATTCGGTGGCCGCGG	2160
Db	2101	GGCAGCTCGGACCTTTACTTGTCTCAGGAGCATGCGATGTCTATTCGGTGGCCGCGG	2160
QY	2161	GGCGACAGCAGGGGAGCGCTACTTCCCGCCAGCGCCGCTCTCTACTTTGAAGGCTCTTTCG	2220
Db	2161	GGCGACAGCAGGGGAGCGCTACTTCCCGCCAGCGCCGCTCTCTACTTTGAAGGCTCTTTCG	2220
QY	2221	GGCGGTCACCTGTCTGCCCCCTCGGGGACAGGTGTGGGCACTTTTCGGGCTGCGGTGTC	2280
Db	2221	GGCGGTCACCTGTCTGCCCCCTCGGGGACAGGTGTGGGCACTTTTCGGGCTGCGGTGTC	2280
QY	2281	ACCCGAGGGGTTGCGAAGCGGTGACCTTTGTACCGCTCGAGTCTATGTGAAGAACCATATG	2340
Db	2281	ACCCGAGGGGTTGCGAAGCGGTGACCTTTGTACCGCTCGAGTCTATGTGAAGAACCATATG	2340
QY	2341	CGGTCCCCCGGTCTTTCACGGACAACTCGTCCCCCTCCGGCCGTACCGCAGACATTTCCAGGTG	2400
Db	2341	CGGTCCCCCGGTCTTTCACGGACAACTCGTCCCCCTCCGGCCGTACCGCAGACATTTCCAGGTG	2400
QY	2401	GCCCATCTACACGCCCTTACTGCTAGCGCAAGAGCACTAAAGTGCCTGCGGTATGCA	2460
Db	2401	GCCCATCTACACGCCCTTACTGCTAGCGCAAGAGCACTAAAGTGCCTGCGGTATGCA	2460
QY	2461	GCCAAGGATTAAGGTGCTTCTGCTGAACCGCTCGCTCGCGCCACCTAGGTTTCGGG	2520
Db	2461	GCCAAGGATTAAGGTGCTTCTGCTGAACCGCTCGCTCGCGCCACCTAGGTTTCGGG	2520
QY	2521	GGGTATATCTTAAGGACATGCTATCGACCCCTTAAACATCAGAACCCGGGTAAAGCAATC	2580
Db	2521	GGGTATATCTTAAGGACATGCTATCGACCCCTTAAACATCAGAACCCGGGTAAAGCAATC	2580
QY	2581	ACACGGGTGCCCCCATACGTACTCCACCTATGGAAGGTTCTTTCGCCACGGGTGCTGC	2640
Db	2581	ACACGGGTGCCCCCATACGTACTCCACCTATGGAAGGTTCTTTCGCCACGGGTGCTGC	2640
QY	2641	TCTGGGGGCGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT	2700
Db	2641	TCTGGGGGCGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT	2700
QY	2701	ATCCTGGGCATCGGCACAGTCTCTGGACCAAGGGAGACGGCTGGAGCGGACTCGTCGTG	2760
Db	2701	ATCCTGGGCATCGGCACAGTCTCTGGACCAAGGGAGACGGCTGGAGCGGACTCGTCGTG	2760
QY	2761	CTCGCCACCGCTACGGCTCCGGGATCGGTCACTCGTGCACATCCAAACATCGAGGAGTG	2820
Db	2761	CTCGCCACCGCTACGGCTCCGGGATCGGTCACTCGTGCACATCCAAACATCGAGGAGTG	2820
QY	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880
Db	2821	GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880
QY	2881	AAGGGGGGAGGACCTCACTTTTCTGCCATTCGAAGAAATGTGATGAGTTCGCGCG	2940
Db	2881	AAGGGGGGAGGACCTCACTTTTCTGCCATTCGAAGAAATGTGATGAGTTCGCGCG	2940
QY	2941	AAGCTTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGCGCTTCAATGATCCGTC	3000
Db	2941	AAGCTTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGCGCTTCAATGATCCGTC	3000
QY	3001	ATACCAACTAGCGGAGACGCTCATTTGCTGTAGCAACGGACGCTCTAATAGCGGCTTTACC	3060
Db	3001	ATACCAACTAGCGGAGACGCTCATTTGCTGTAGCAACGGACGCTCTAATAGCGGCTTTACC	3060
QY	3061	GGCGATTTCCGACTCAGTGTGATCGACTGCAATACATGTGTCCACCCAGACAGTTCGACTCAGC	3120
Db	3061	GGCGATTTCCGACTCAGTGTGATCGACTGCAATACATGTGTCCACCCAGACAGTTCGACTCAGC	3120
QY	3121	CTGGACCCGACCTTACCAATTTGAGACGACGACGCTGCCCAAGACCGGCTGTACGCTCG	3180
Db	3121	CTGGACCCGACCTTACCAATTTGAGACGACGACGCTGCCCAAGACCGGCTGTACGCTCG	3180

QY 3181 CAGCGGAGGACGACTGTTAGGGGACGAGTGGCAATTACAGTGTGTGACTCCAGGA 3240  
DB |||||  
QY 3191 CAGCGGAGGACGAGACTGGTAGGGGACGAGTGGGCAATTACAGTGTGTGACTCCAGGA 3240  
DB |||||  
QY 3241 GAACGGCCCTCGGGCATGTTCCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
DB |||||  
QY 3241 GAACGGCCCTCGGGCATGTTCCGATTCCTCGGTTCTGTGCGAGTGTATGACGCGGGCTGT 3300  
DB |||||  
QY 3301 GCTTGTAGAGAGTCAACGCCCGCGAGACCTCAGTTAGTGTGCGGGCTTACTAAACACA 3360  
DB |||||  
QY 3301 GCTTGTAGAGAGTCAACGCCCGCGAGACCTCAGTTAGTGTGCGGGCTTACTAAACACA 3360  
DB |||||  
QY 3361 CCAGGTTGCCGCTCCAGGACCATCTGGAGTTCCTGGGAGAGCTTTTACAGGCTC 3420  
DB |||||  
QY 3361 CCAGGTTGCCGCTCCAGGACCATCTGGAGTTCCTGGGAGAGCTTTTACAGGCTC 3420  
DB |||||  
QY 3421 ACCACATAGACGCCCAATTTCTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTAC 3480  
DB |||||  
QY 3421 ACCACATAGACGCCCAATTTCTGTCCAGACTAAGCAGGAGGAGACAACTTCCCTAC 3480  
DB |||||  
QY 3481 CTGTAGCATACAGGCTACCGGTGTGCGGCGAGGCTCAGGCTCCACTCCATCGTGGAC 3540  
DB |||||  
QY 3481 CTGTAGCATACAGGCTACCGGTGTGCGGCGAGGCTCAGGCTCCACTCCATCGTGGAC 3540  
DB |||||  
QY 3541 CAAATGTGGAAGTGTCTCATACGCTAAGCTCAGCTGACGCGGCGCAACGCCCTGCTG 3600  
DB |||||  
QY 3541 CAAATGTGGAAGTGTCTCATACGCTAAGCTCAGCTGACGCGGCGCAACGCCCTGCTG 3600  
DB |||||  
QY 3601 TATAGGCTGGAGCGGTTCAAACGAGGTTACTACACACACCCCATAAACAAATACATC 3660  
DB |||||  
QY 3601 TATAGGCTGGAGCGGTTCAAACGAGGTTACTACACACACCCCATAAACAAATACATC 3660  
DB |||||  
QY 3661 ATGCAATGATGTGCGGTGACCTGGAGTGTCTACAGGACACTGGTGTCTGGTAGGCGGA 3720  
DB |||||  
QY 3661 ATGCAATGATGTGCGGTGACCTGGAGTGTCTACAGGACACTGGTGTCTGGTAGGCGGA 3720  
DB |||||  
QY 3721 GTCTAGCAGCTCTGGCGCGGTATGCTGTGACACAGCAGCGGTGCTATTTGGGCGAG 3780  
DB |||||  
QY 3721 GTCTAGCAGCTCTGGCGCGGTATGCTGTGACACAGCAGCGGTGCTATTTGGGCGAG 3780  
DB |||||  
QY 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTACCGGGAGTTC 3840  
DB |||||  
QY 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTCGCGACAGGGAAGTCTTTACCGGGAGTTC 3840  
DB |||||  
QY 3841 GATGAGATGGAAGTGGCGCTCAACCTCCCTTTACATCGAAACAGGGAATGCAGCTCGCC 3900  
DB |||||  
QY 3901 GAACAAATTCACACAGAGGCAATCGGGTGTGCAACAGCAGCCACACAGCAGCGAGGCT 3960  
DB |||||  
QY 3901 GAACAAATTCACACAGAGGCAATCGGGTGTGCAACAGCAGCCACACAGCAGCGAGGCT 3960  
DB |||||  
QY 3961 GCTGCTCCCGTGTGGAAATCCAAAGTGGCGGACCTCGAAGCTTCTGGGCGAAGCATATG 4020  
DB |||||  
QY 3961 GCTGCTCCCGTGTGGAAATCCAAAGTGGCGGACCTCGAAGCTTCTGGGCGAAGCATATG 4020  
DB |||||  
QY 4021 TGGAAATTTATACCGGGAATACAAATTTAGCAGGCTGTCTCACTCTGCTCGGCAACCCC 4080  
DB |||||  
QY 4021 TGGAAATTTATACCGGGAATACAAATTTAGCAGGCTGTCTCACTCTGCTCGGCAACCCC 4080  
DB |||||  
QY 4081 GCGATAGCATCACTGATGGGCAATTCACAGCTTCTATACAGCCCGCTACACCCCAACAT 4140  
DB |||||  
QY 4081 GCGATAGCATCACTGATGGGCAATTCACAGCTTCTATACAGCCCGCTACACCCCAACAT 4140  
DB |||||  
QY 4141 ACCCTCTGTTTAAACATCTCTGGGGGATGGGTGGCGCGCCCAACTTGTCTCTCCAGGCT 4200  
DB |||||  
QY 4141 ACCCTCTGTTTAAACATCTCTGGGGGATGGGTGGCGCGCCCAACTTGTCTCTCCAGGCT 4200  
DB |||||  
QY 4201 GCTTCTGCTTCTGTAGGCGCGGATCGCTGAGCGGCTGTGTGGCAGCATAGGCTTGGG 4260  
DB |||||  
QY 4201 GCTTCTGCTTCTGTAGGCGCGGATCGCTGAGCGGCTGTGTGGCAGCATAGGCTTGGG 4260  
DB |||||

QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGGACAGGCGGCTCGTGCC 4320  
DB |||||  
QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGGACAGGCGGCTCGTGCC 4320  
DB |||||  
QY 4321 TTTAAGGTTCATGAGCGGAGATGCTCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
DB |||||  
QY 4321 TTTAAGGTTCATGAGCGGAGATGCTCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
DB |||||  
QY 4381 ATCTCTCTCCCTGGCGCCCTAGTCTCGGGTCTGTGTCGAGCGATATCTCGTCTGGCAC 4440  
DB |||||  
QY 4381 ATCTCTCTCCCTGGCGCCCTAGTCTCGGGTCTGTGTCGAGCGATATCTCGTCTGGCAC 4440  
DB |||||  
QY 4441 GTGGGCGGAGGAGGCGCTGTGAGTGTGATGAACCGGCTGATAGCTTGTCTTGGCG 4500  
DB |||||  
QY 4441 GTGGGCGGAGGAGGCGCTGTGAGTGTGATGAACCGGCTGATAGCTTGTCTTGGCG 4500  
DB |||||  
QY 4501 GGTAAACACAGCTCTCCCCCAGCCTATGTGTGCTCAGAGCGAGCTGACAGCGTGTCACT 4560  
DB |||||  
QY 4501 GGTAAACACAGCTCTCCCCCAGCCTATGTGTGCTCAGAGCGAGCTGACAGCGTGTCACT 4560  
DB |||||  
QY 4561 CAGATCTCTCTAGTCTTTCATCACTCAGCTGTCTGAAGGCTTCCACAGTGGATCAAC 4620  
DB |||||  
QY 4561 CAGATCTCTCTAGTCTTTCATCACTCAGCTGTCTGAAGGCTTCCACAGTGGATCAAC 4620  
DB |||||  
QY 4621 GAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTGAGAGATGTTTGGGATTTGATATGC 4680  
DB |||||  
QY 4621 GAGGACTGCTCCACGCCATGCTCCGGCTCGTGGCTGAGAGATGTTTGGGATTTGATATGC 4680  
DB |||||  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCCAGTCTGCGCGGATTCGCGGA 4740  
DB |||||  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGCTCCAGTCCAGTCTGCGCGGATTCGCGGA 4740  
DB |||||  
QY 4741 GTCCCTCTCTCTCATGTCAACGTTGGGTACAAGGAGTCTGCGCGGCGAGCGCATATG 4800  
DB |||||  
QY 4741 GTCCCTCTCTCTCATGTCAACGTTGGGTACAAGGAGTCTGCGCGGCGAGCGCATATG 4800  
DB |||||  
QY 4801 CAAACACAGCTGCTCCATGTGAGGACAGATCAACCGGACATGTGAAAAACGGTTCATGAGG 4860  
DB |||||  
QY 4801 CAAACACAGCTGCTCCATGTGAGGACAGATCAACCGGACATGTGAAAAACGGTTCATGAGG 4860  
DB |||||  
QY 4861 ATCGTGGGCTTAGGACCTGTGATTAACAGTGGCATGGAATTTTCTAGGGGCTGTGCGGGTG 4920  
DB |||||  
QY 4861 ATCGTGGGCTTAGGACCTGTGATTAACAGTGGCATGGAATTTTCTAGGGGCTGTGCGGGTG 4920  
DB |||||  
QY 4921 ACCACGGGCGGCTGACGCGCTCCCGCGCGCAAAATTTATCTAGGGGCTGTGCGGGTG 4980  
DB |||||  
QY 4921 ACCACGGGCGGCTGACGCGCTCCCGCGCGCAAAATTTATCTAGGGGCTGTGCGGGTG 4980  
DB |||||  
QY 4981 GCTGCTGAGGATGAGTGGAGGTTTACCGGGTGGGGATTTTCCACTAGTGAACGGGATG 5040  
DB |||||  
QY 4981 GCTGCTGAGGATGAGTGGAGGTTTACCGGGTGGGGATTTTCCACTAGTGAACGGGATG 5040  
DB |||||  
QY 5041 ACCACTCACAAAGTAAAGTCCCGGTGTCAGTTCGGGCGGCGCAATTTCTTACAGAGATG 5100  
DB |||||  
QY 5041 ACCACTCACAAAGTAAAGTCCCGGTGTCAGTTCGGGCGGCGCAATTTCTTACAGAGATG 5100  
DB |||||  
QY 5101 GATGGGTGGGTGTGACAGTACGCTCCAGCGTGAACCCCTCTTACGGGAGGAGTTC 5160  
DB |||||  
QY 5101 GATGGGTGGGTGTGACAGTACGCTCCAGCGTGAACCCCTCTTACGGGAGGAGTTC 5160  
DB |||||  
QY 5161 ACATTCCTGTGCGGCTCAATCAATCACTGCTCAGCGAGCGGCTCAAGTTCCTATCGGAGCGGAA 5220  
DB |||||  
QY 5161 ACATTCCTGTGCGGCTCAATCAATCACTGCTCAGCGAGCGGCTCAAGTTCCTATCGGAGCGGAA 5220  
DB |||||  
QY 5221 CCGGAGCTAGAGTGTCTTCTCATGCTCAGCGAGCGGCTCCACATTTACAGGAGAGGAGT 5280  
DB |||||  
QY 5221 CCGGAGCTAGAGTGTCTTCTCATGCTCAGCGAGCGGCTCCACATTTACAGGAGAGGAGT 5280  
DB |||||  
QY 5281 GCTAAGGCTAGGCTGGCAGGAGGATCTCCCGCTCTTGGCGAGCTCATCAGCTAGGAG 5340  
DB |||||  
QY 5281 GCTAAGGCTAGGCTGGCAGGAGGATCTCCCGCTCTTGGCGAGCTCATCAGCTAGGAG 5340  
DB |||||  
QY 5341 CTGTCTGCGGCTTCTTGAAGGCAACATGCACTACCGGTCATGACTCCCGGAGCGTGTGAC 5400  
DB |||||



```
QY 7561 AACTGGGCGAGTAAGACCAAGCTCAAACTCACTCAATCCCGGCTGCGTCCAGTTGGAT 7620
Db 7561 AACTGGGCGAGTAAGACCAAGCTCAAACTCACTCAATCCCGGCTGCGTCCAGTTGGAT 7620
QY 7621 TTATCCAGCTGGTTCTGTTGCTGTGTACAGCGGGGAGACATATATACAGAGCTGTCTCGT 7680
Db 7621 TTATCCAGCTGGTTCTGTTGCTGTGTACAGCGGGGAGACATATATACAGAGCTGTCTCGT 7680
QY 7681 GCCCGACCCGCTGGTTCACTGTGGTCCCTACTCTCTAATCTCTGTAGGGGTAGGCACTAT 7740
Db 7681 GCCCGACCCGCTGGTTCACTGTGGTCCCTACTCTCTAATCTCTGTAGGGGTAGGCACTAT 7740
QY 7741 CTACTCTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTGTTTT 7800
Db 7741 CTACTCTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTTT 7800
QY 7801 TTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860
Db 7801 TTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 7860
QY 7861 TTTTTCCTCTTTTTTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 7920
Db 7861 TTTTTCCTCTTTTTTCTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 7920
QY 7921 TAGCTGTGAAGAGTCCGTGAGCGCTTGACTGACAGAGAGTCTGATAGTCCCTCTCTGC 7980
Db 7921 TAGCTGTGAAGAGTCCGTGAGCGCTTGACTGACAGAGAGTCTGATAGTCCCTCTCTGC 7980
QY 7981 AGATCAAGT 7989
Db 7981 AGATCAAGT 7989

RESULT 5
US-10-125-920-1
; Sequence 1, Application US/10125920
; Publication No. US20030176503A1
; GENERAL INFORMATION:
; APPLICANT: Altamura, Sergio
; APPLICANT: Kock, Uwe
; TITLE OF INVENTION: TREATING HEPATITIS C VIRAL INFECTIONS
; TITLE OF INVENTION: WITH THIOSEMICARBAZONE COMPOUNDS
; FILE REFERENCE: IT0010Y
; CURRENT APPLICATION NUMBER: US/10/125,920
; CURRENT FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/285,195
; PRIOR FILING DATE: 2001-04-20
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: Fast-SEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 10690
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Plasmid
US-10-125-920-1

Query Match 100.0%; Score 7987.4; DB 15; Length 10690;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7988; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATTGGGGGAGACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60
Db 1 GCCAGCCCCGATTGGGGGAGACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60
QY 61 TCTTCAGCAGAAAGCGTCTAGCCATGCGCTTAGTATGAGTGTGCTGAGCTCCAGGAC 120
Db 61 TCTTCAGCAGAAAGCGTCTAGCCATGCGCTTAGTATGAGTGTGCTGAGCTCCAGGAC 120
QY 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTCGGGAACCGGTGAGTACACCGGAATGCGAG 180
Db 121 CCCCCCTCCGGGAGAGCCATAGTGGTCTCGGGAACCGGTGAGTACACCGGAATGCGAG 180
```

```
QY 181 GACGACCGGTCCTTTCTTTGGATCAACCCCGCTCAATGCTGAGATTTGGGGTGCCTCC 240
Db 181 GACGACCGGTCCTTTCTTTGGATCAACCCCGCTCAATGCTGAGATTTGGGGTGCCTCC 240
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGTCCGAAAGCCCTTGTGGTACTGCTGATAGG 300
Db 241 GCGAGACTGTAGCCGAGTAGTGTGGTCCGAAAGCCCTTGTGGTACTGCTGATAGG 300
QY 301 GTGCTTGGAGTGCCTCCGGAGGTCCTGTAGACCGTGCACCATGAGCAAGCAATCCTAAAC 360
Db 301 GTGCTTGGAGTGCCTCCGGAGGTCCTGTAGACCGTGCACCATGAGCAAGCAATCCTAAAC 360
QY 361 CTCAAGAAAACCAAGAGGCGCCCATGATTGAACAAGATGGATGGACGAGGTTCTC 420
Db 361 CTCAAGAAAACCAAGAGGCGCCCATGATTGAACAAGATGGATGGACGAGGTTCTC 420
QY 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAGACAATCGGCTGCT 480
Db 421 CGGCGCTTGGGTGAGAGGCTATTTCGGCTATGACTGGGCAACAGACAATCGGCTGCT 480
QY 481 CTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGGCCCGGTTCTTTTGTCAAGACCG 540
Db 481 CTGATGCGCGCTGTTCCGGCTGTACGCGAGGGCGGCCCGGTTCTTTTGTCAAGACCG 540
QY 541 ACCTGTCCGTGCCCTGAAATGAATCTGAGAGAGAGGCGCGGCTATCGTGGCTGGCCA 600
Db 541 ACCTGTCCGTGCCCTGAAATGAATCTGAGAGAGAGGCGCGGCTATCGTGGCTGGCCA 600
QY 601 CGACGGCGCTTCTTTCGCGAGCTGTCTCGACGTTGTCACTGAAGCGGAGGAGACTGCG 660
Db 601 CGACGGCGCTTCTTTCGCGAGCTGTCTCGACGTTGTCACTGAAGCGGAGGAGACTGCG 660
QY 661 TGCTATTGGGCGAGTGCCTGGGCGAGGATCTCTGTGCTATCTCACTTGTCTCTCCGAGA 720
Db 661 TGCTATTGGGCGAGTGCCTGGGCGAGGATCTCTGTGCTATCTCACTTGTCTCTCCGAGA 720
QY 721 AAGTATCCATGCTGATGCAATGCGGCGCTGCTAGCTGTGATCGGCTACCTGCTG 780
Db 721 AAGTATCCATGCTGATGCAATGCGGCGCTGCTAGCTGTGATCGGCTACCTGCTG 780
QY 781 CATTCGACACCAAGCGAATCATCGCATCGAGCGAGCACTGCTGATCGAGCGGCTG 840
Db 781 CATTCGACACCAAGCGAATCATCGCATCGAGCGAGCACTGCTGATCGAGCGGCTG 840
QY 841 TTGTCGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTG 900
Db 841 TTGTCGATCAGGATGATCTGGAAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTG 900
QY 901 CCAGGCTCAAGGCGCGCATGCCGAGCGGAGGATCTCGTGTGACCATGCGGATGCT 960
Db 901 CCAGGCTCAAGGCGCGCATGCCGAGCGGAGGATCTCGTGTGACCATGCGGATGCT 960
QY 961 GCTTGGCGAATATCATGTTGAAATGCGCTTTCTGATTCATGATGCTGCGGCG 1020
Db 961 GCTTGGCGAATATCATGTTGAAATGCGCTTTCTGATTCATGATGCTGCGGCG 1020
QY 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGGTTGGTACCCGCTGATTTGCTGAAGAGC 1080
Db 1021 TGGGTGTGGCGGACCGCTATCAGGACATAGGTTGGTACCCGCTGATTTGCTGAAGAGC 1080
QY 1081 TTGGCGGAAATGGGCTGACCGCTTCTGCTGTTTACGGTATCGCGCTCCCGATTCG 1140
Db 1081 TTGGCGGAAATGGGCTGACCGCTTCTGCTGTTTACGGTATCGCGCTCCCGATTCG 1140
QY 1141 AGCGCATCGGCTTCTGAGAGGTTCTTGAAGTCTTCTGAGTTTAAACAGACCAAGC 1200
Db 1141 AGCGCATCGGCTTCTGAGAGGTTCTTGAAGTCTTCTGAGTTTAAACAGACCAAGC 1200
QY 1201 GTTTCCTCTAGCGGAGTCAATTCGCGCTTCTCCCTCCCGCTCCCGCTTACTTGGC 1260
Db 1201 GTTTCCTCTAGCGGAGTCAATTCGCGCTTCTCCCTCCCGCTCCCGCTTACTTGGC 1260
```

1261 CGAAGCCGCTTGGAAATAAGCCGGTGTGGTGGTTGTCTATATGTTATTTTCCACCATATTG 1320  
1261 CGAAGCCGCTTGGAAATAAGCCGGTGTGGTGGTTGTCTATATGTTATTTTCCACCATATTG 1320  
1321 CGGCTCTTTTGGCAATGTGAGGGCCCGAAACCTTGGCCCTGTCTTCTTTCAGCGACATTCCT 1380  
1321 CGGCTCTTTTGGCAATGTGAGGGCCCGAAACCTTGGCCCTGTCTTCTTTCAGCGACATTCCT 1380  
1381 AGGGGCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGCTGTGAAGGAGCA 1440  
1381 AGGGGCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGCTGTGAAGGAGCA 1440  
1441 GTTCTCTCGAAGCTTTTGAACAAACAAAGGTCTGTAGCAGCCCTTTTGCAGGACGCG 1500  
1441 GTTCTCTCGAAGCTTTTGAACAAACAAAGGTCTGTAGCAGCCCTTTTGCAGGACGCG 1500  
1501 AACCCGCCCACTTGGGACAGGTGCTCTCGGCCAAAGCCACGCTGTATAAGATACACCT 1560  
1501 AACCCGCCCACTTGGGACAGGTGCTCTCGGCCAAAGCCACGCTGTATAAGATACACCT 1560  
1561 GCAAGGCGGCAACACCCAGTGCACGTTGTGAGTTGGATAGTTCTGGAAGAGTCAAA 1620  
1561 GCAAGGCGGCAACACCCAGTGCACGTTGTGAGTTGGATAGTTCTGGAAGAGTCAAA 1620  
1621 TGGCTCTCTCAAGCTTATCAACAGGGGCTGAAGATGCCCCAGAGGTACCCCATTTGT 1680  
1621 TGGCTCTCTCAAGCTTATCAACAGGGGCTGAAGATGCCCCAGAGGTACCCCATTTGT 1680  
1681 ATGGGATCTGATCTGGGCGCTCGTGCACATGCTTTACATGTTTGTAGTTCGAGTTAAAA 1740  
1681 ATGGGATCTGATCTGGGCGCTCGTGCACATGCTTTACATGTTTGTAGTTCGAGTTAAAA 1740  
1741 AAGCTTAGGCCCGCCGAAACACCGGGAAGTGTGTTTCTTTGAAAAACAGATATAC 1800  
1741 AAGCTTAGGCCCGCCGAAACACCGGGAAGTGTGTTTCTTTGAAAAACAGATATAC 1800  
1801 ATGGCGCTATTACGCTTACTCCCAACAGCGAGGCTACTTGGCTGCATCATCACT 1860  
1801 ATGGCGCTATTACGCGCTACTCCCAACAGCGAGGCTACTTGGCTGCATCATCACT 1860  
1861 AGCTCACAGCGCGGACAGAAACAGGTGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
1861 AGCTCACAGCGCGGACAGAAACAGGTGAGGGGAGGTCCAAAGTGTCTCCACCGCA 1920  
1921 ACACAATCTTTCTGGCGAAGCTGGTCAATGGCGTGTGTTGAGCTGTCTATCATGTGTC 1980  
1921 ACACAATCTTTCTGGCGAAGCTGGTCAATGGCGTGTGTTGAGCTGTCTATCATGTGTC 1980  
1981 GGCTCAAGACCTTGCAGCGCCAAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
1981 GGCTCAAGACCTTGCAGCGCCAAAGGGCCCAATCACCCAAATGTACCAATGTGGAC 2040  
2041 CAGGACCTCGTGGTGGCAAGCGCCCGCGGGCGGTTCTTTGACCAATGCAACCTGC 2100  
2041 CAGGACCTCGTGGTGGCAAGCGCCCGCGGGCGGTTCTTTGACCAATGCAACCTGC 2100  
2101 GGAGCTCGACCTTTACTTGGTCAAGAGGATGCGGATGCTATTCGGGTGCGCGCGG 2160  
2101 GGAGCTCGACCTTTACTTGGTCAAGAGGATGCGGATGCTATTCGGGTGCGCGCGG 2160  
2161 GGCGACAGAGGGGAGCCTTACTCTCCCGAGGCGGCTCTTACTTGAAGGGCTCTTCG 2220  
2161 GGCGACAGAGGGGAGCCTTACTCTCCCGAGGCGGCTCTTACTTGAAGGGCTCTTCG 2220  
2221 GGCGGTCACCTGTCTGCCCTCGGGGACAGCTGTGGGCAATCTTTGGGCTGCGGTGTC 2280  
2221 GGCGGTCACCTGTCTGCCCTCGGGGACAGCTGTGGGCAATCTTTGGGCTGCGGTGTC 2280  
2281 ACCGAGGGGTGGAGGGGCTGGAATTTGTACCGCTCGAGTCTATGAAACCACTATG 2340  
2281 ACCGAGGGGTGGAGGGGCTGGAATTTGTACCGCTCGAGTCTATGAAACCACTATG 2340  
2341 CGGTCCCGGCTTTCACGGCAAACTCGTCCCGCTCGGCGGTACCGCAGACATTCAGGTG 2400

2341 CGGTCCCGGCTTTCACGGCAAACTCGTCCCGCGGTACCGCAGACATTCAGGTG 2400  
2401 GCCCATCTACAGCCCTCTACTGTAGCGCAGAGCACTAAGGTGCCGCTCGTATGCA 2460  
2401 GCCCATCTACAGCCCTCTACTGTAGCGCAGAGCACTAAGGTGCCGCTCGTATGCA 2460  
2461 GCCAAGGGTATAAGGTCTTCTGTAACCCGTCGTCGCCGCCACCTAGTGTTCGGG 2520  
2461 GCCAAGGGTATAAGGTCTTCTGTAACCCGTCGTCGCCGCCACCTAGTGTTCGGG 2520  
2521 GGGTATATGTTCTAAGGCACATGGTATCGACCCCTAAACATCAGAAACCGGGTAAAGCAATC 2580  
2521 GGGTATATGTTCTAAGGCACATGGTATCGACCCCTAAACATCAGAAACCGGGTAAAGCAATC 2580  
2581 ACACGGGTGCCCCCATCAGCTACTCCACTATGGAAGTCTTTTCCGCCAGGTGTGTC 2640  
2581 ACACGGGTGCCCCCATCAGCTACTCCACTATGGAAGTCTTTTCCGCCAGGTGTGTC 2640  
2641 TCTGGGGCGGCTATGACATCATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
2641 TCTGGGGCGGCTATGACATCATATATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
2701 ATCTGGGCGATCGGCACACTCCTGACACMAAGCGGAGACGGCTGGAGCGGACTTCGTGTC 2760  
2701 ATCTGGGCGATCGGCACACTCCTGACACMAAGCGGAGACGGCTGGAGCGGACTTCGTGTC 2760  
2761 CTGCGCACCGCTACCGCTCGGGATCGGTACCGTGCCACATCCAAAGCTCGAGGAGGTG 2820  
2761 CTGCGCACCGCTACCGCTCGGGATCGGTACCGTGCCACATCCAAAGCTCGAGGAGGTG 2820  
2821 GCTCTCTCAGCACTCGGAAATCCCTTTTATGGCAAGGCCATCCCTCCAGAGCAATC 2880  
2821 GCTCTCTCAGCACTCGGAAATCCCTTTTATGGCAAGGCCATCCCTCCAGAGCAATC 2880  
2881 AAGGGGGGAGGACCTCTCTTTTCTGCCATTCAGAGAAATGTGATGCTCGCGCGG 2940  
2881 AAGGGGGGAGGACCTCTCTTTTCTGCCATTCAGAGAAATGTGATGCTCGCGCGG 2940  
2941 AAGCTCTCGGCTCGGACTCAATGCTGTAGCATATTCGGGGCTTGATGTATCGGTC 3000  
2941 AAGCTCTCGGCTCGGACTCAATGCTGTAGCATATTCGGGGCTTGATGTATCGGTC 3000  
3001 ATACCAACTAGCGGAGACGCTATTGTCTAGCAACCGAGCGCTCTAATGACGGCTTAC 3060  
3001 ATACCAACTAGCGGAGACGCTATTGTCTAGCAACCGAGCGCTCTAATGACGGCTTAC 3060  
3061 GGGATTTTCGACTCAGTGATCGACTGCAATATCATGTGTCAACCCAGACAGTTCGACTTCAGC 3120  
3061 GGGATTTTCGACTCAGTGATCGACTGCAATATCATGTGTCAACCCAGACAGTTCGACTTCAGC 3120  
3121 CTGGACCGGACCTTACCATTTGAGACGACCGTGCACAAAGCGGGTGTCAAGCTCG 3180  
3121 CTGGACCGGACCTTACCATTTGAGACGACCGTGCACAAAGCGGGTGTCAAGCTCG 3180  
3181 CAGCGCGAGGAGGAGTGTAGGGCAGGATGCGGCAATTTACAGTTTGTGATCTCCAGGA 3240  
3181 CAGCGCGAGGAGGAGTGTAGGGCAGGATGCGGCAATTTACAGTTTGTGATCTCCAGGA 3240  
3241 GAAACGGCCCTCGGGCATGTTTCGATTCCTCGGTTCTGTGCGAGTGTCTATGACGGGGCTGT 3300  
3241 GAAACGGCCCTCGGGCATGTTTCGATTCCTCGGTTCTGTGCGAGTGTCTATGACGGGGCTGT 3300  
3301 GCTTGTACGAGCTACGGCCCGGAGACTCTAGTTAGGTTCGGGCTTACCTTAAACACA 3360  
3301 GCTTGTACGAGCTACGGCCCGGAGACTCTAGTTAGGTTCGGGCTTACCTTAAACACA 3360  
3361 CCAGGGTTGCCCTCTGCGCAGGACCTCTAGTTAGGTTCGGGCTTACCTTAAACACA 3420  
3361 CCAGGGTTGCCCTCTGCGCAGGACCTCTAGTTAGGTTCGGGCTTACCTTAAACACA 3420  
3421 ACCCAGATAGACGCCCATTTCTGTGTCAGTCTAAGCAGGAGGAGCAACTTCCCTTAC 3480



Db	3421	ACCCACATAGACGCCCAATTTCTTGTCACAGACTAAGCAGGACGAGACAACATTTCCCTAC	3480
Qy	3481	CTGGTAGCATACAGGCTACGGTGTGGCCAGGGCTCAGGCTCCAGCTCCATCGTGGGAC	3540
Db	3481	CTGGTAGCATACAGGCTACGGTGTGGCCAGGGCTCAGGCTCCATCGTGGGAC	3540
Qy	3541	CAAAATGTGGAAGTGTCTCATACGGCTAAAGCCTAAGCTGCACGGGCAACGCCCTGCTG	3600
Db	3541	CAAAATGTGGAAGTGTCTCATACGGCTAAAGCCTAAGCTGCACGGGCAACGCCCTGCTG	3600
Qy	3601	TATAGGCTGGGAGCCGTTCAAAACAGAGTTACTACACACGCCCATAAACCAATACATC	3660
Db	3601	TATAGGCTGGGAGCCGTTCAAAACAGAGTTACTACACACGCCCATAAACCAATACATC	3660
Qy	3661	ATGGCATCATGTCCGGCTGACCTCGAGTTCGTCACGAGCAGCTGGTGTGATGGCCGA	3720
Db	3661	ATGGCATCATGTCCGGCTGACCTCGAGTTCGTCACGAGCAGCTGGTGTGATGGCCGA	3720
Qy	3721	GTCTAGCAGCTCTGGCCGCGTATTGCTGACAAACAGGACGCTGCTATTGTGGCCAGG	3780
Db	3721	GTCTAGCAGCTCTGGCCGCGTATTGCTGACAAACAGGACGCTGCTATTGTGGCCAGG	3780
Qy	3781	ATCATCTTGTCCGAAAGCCGGCCATCATTTCCGACAGGGAAGTCTTTACCGGAGTTC	3840
Db	3781	ATCATCTTGTCCGAAAGCCGGCCATCATTTCCGACAGGGAAGTCTTTACCGGAGTTC	3840
Qy	3841	GATGAGATGGAAGAGTGGCCCTCACACCTCCCTTACATCGACAGGGAATGAGCTGCC	3900
Db	3841	GATGAGATGGAAGAGTGGCCCTCACACCTCCCTTACATCGACAGGGAATGAGCTGCC	3900
Qy	3901	GAAACATTTCAAACAGAGGAATCGGTTGTGCAAAACAGCCACCAAGCAAGCGAGGCT	3960
Db	3901	GAAACATTTCAAACAGAGGAATCGGTTGTGCAAAACAGCCACCAAGCAAGCGAGGCT	3960
Qy	3961	GCTCTCCCGTGGTAATCAAGTGGGGAGCCCTCGAAGCCTTCTGGGCGAAGCATATG	4020
Db	3961	GCTCTCCCGTGGTAATCAAGTGGGGAGCCCTCGAAGCCTTCTGGGCGAAGCATATG	4020
Qy	4021	TGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTGCTGCAACCCC	4080
Db	4021	TGGAATTTTCATCAGCGGATACAATATTTAGCAGGCTTGTCCACTCTGCTGCAACCCC	4080
Qy	4081	CGGATAGCATCACTGATGGCAATTCACAGCCTTATCACAGCCCGCTCACCAACCAACAT	4140
Db	4081	CGGATAGCATCACTGATGGCAATTCACAGCCTTATCACAGCCCGCTCACCAACCAACAT	4140
Qy	4141	ACCTCTCTGTTTACATCTCTGGGGGATGGTGGCCGCCCAACTTGTCTCTCCAGCGCT	4200
Db	4141	ACCTCTCTGTTTACATCTCTGGGGGATGGTGGCCGCCCAACTTGTCTCTCCAGCGCT	4200
Qy	4201	GCTTCTCTCTTGTAGGCGCGGCATCGCTGAGCGGCTGTGTCAGCATAGGCTTTGGG	4260
Db	4201	GCTTCTCTCTTGTAGGCGCGGCATCGCTGAGCGGCTGTGTCAGCATAGGCTTTGGG	4260
Qy	4261	AAGTGCTGTGGATATTTTGGCAGGTTATGGAGAGGGGTGGCAGGCGGCTCGTGGCC	4320
Db	4261	AAGTGCTGTGGATATTTTGGCAGGTTATGGAGAGGGGTGGCAGGCGGCTCGTGGCC	4320
Qy	4321	TTTAAAGTTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTTCCCTGCT	4380
Db	4321	TTTAAAGTTCATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTTCCCTGCT	4380
Qy	4381	ATCTCTCTCTCTGCGCCCTAGTCTGCGGGTGTGTCGCGAGGATGCTGCGGCAC	4440
Db	4381	ATCTCTCTCTCTGCGCCCTAGTCTGCGGGTGTGTCGCGAGGATGCTGCGGCAC	4440
Qy	4441	GTGGGCCAGGGAGGGGGTGTGTCAGTGGATGAACCGGCTGTAGCGTTCTGCTTGGCG	4500
Db	4441	GTGGGCCAGGGAGGGGGTGTGTCAGTGGATGAACCGGCTGTAGCGTTCTGCTTGGCG	4500
Qy	4501	GGTAAACACGCTCTCCCGCCACGCACTATGTCCTGAGAGCAGCGCTGACGACGCTGCT	4560
Db	4501	GGTAAACACGCTCTCCCGCCACGCACTATGTCCTGAGAGCAGCGCTGACGACGCTGCT	4560
Qy	4561	CAGATCTCTCTAGTCTTTACATCATCTAGCTGTGTAAGAGAGGCTTCAACAGTGGATCAAC	4620
Db	4561	CAGATCTCTCTAGTCTTTACATCATCTAGCTGTGTAAGAGAGGCTTCAACAGTGGATCAAC	4620
Qy	4621	GAGGACTCTCCACGCCATGCTCCGGCTCGTGGCTTAAGAGATGTTTGGGATTTGGATATGC	4680
Db	4621	GAGGACTCTCCACGCCATGCTCCGGCTCGTGGCTTAAGAGATGTTTGGGATTTGGATATGC	4680
Qy	4681	ACGGTGTGACTGATTTTCAAGACCTTGGCTCCAGTCCAAAGCTCTCTCCCGGATTTGGCGGA	4740
Db	4681	ACGGTGTGACTGATTTTCAAGACCTTGGCTCCAGTCCAAAGCTCTCTCCCGGATTTGGCGGA	4740
Qy	4741	GTCCCTCTCTCTCATGTCAACGGTGTGCAAGGAGTCTGGCGGGCGACGGCATCATG	4800
Db	4741	GTCCCTCTCTCTCATGTCAACGGTGTGCAAGGAGTCTGGCGGGCGACGGCATCATG	4800
Qy	4801	CAAAACCACTCCCATGTTGGAGCACAGATCAACGGACATGTGAAACACGGTTCCATGAGG	4860
Db	4801	CAAAACCACTCCCATGTTGGAGCACAGATCAACGGACATGTGAAACACGGTTCCATGAGG	4860
Qy	4861	ATCGTGGGCTTAGGACCTGTAGTAACACAGTGGCATGGAACATTTCCCATTAACCGCTAC	4920
Db	4861	ATCGTGGGCTTAGGACCTGTAGTAACACAGTGGCATGGAACATTTCCCATTAACCGCTAC	4920
Qy	4921	ACACGGGCTCTGACGCCCTCCCGCGCAAAATTTCTAGGGCGCTGTGGCGGGTG	4980
Db	4921	ACACGGGCTCTGACGCCCTCCCGCGCAAAATTTCTAGGGCGCTGTGGCGGGTG	4980
Qy	4981	GCTCTGAGGAGTAGTGGAGTTACCGGGTGGGGATTTCCACTAGTACGGGCTG	5040
Db	4981	GCTCTGAGGAGTAGTGGAGTTACCGGGTGGGGATTTCCACTAGTACGGGCTG	5040
Qy	5041	ACCACTGACAAAGTAAAGTGGCGCTGTCAGGTTCCGGCCCCCGCAATTTCTCACAGAACTG	5100
Db	5041	ACCACTGACAAAGTAAAGTGGCGCTGTCAGGTTCCGGCCCCCGCAATTTCTCACAGAACTG	5100
Qy	5101	GATGGGTGGGTTGCACAGTAGCTCCAGCGTGCAAAACCCCTCTTACGGGAGGAGTGC	5160
Db	5101	GATGGGTGGGTTGCACAGTAGCTCCAGCGTGCAAAACCCCTCTTACGGGAGGAGTGC	5160
Qy	5161	ACATTCCTGTCGGCTCAATCAATCACTGTTGGGTTCAGAGTCCCATGCGAGCGCGAA	5220
Db	5161	ACATTCCTGTCGGCTCAATCAATCACTGTTGGGTTCAGAGTCCCATGCGAGCGCGAA	5220
Qy	5221	CCGACCTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACGGCGGAGACG	5280
Db	5221	CCGACCTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTTACGGCGGAGACG	5280
Qy	5281	GCTAAGCTAGGCTGGCGAGGGATCTCCCTCTCTTGGCCAGCTCATCAGCTAGCCAG	5340
Db	5281	GCTAAGCTAGGCTGGCGAGGGATCTCCCTCTCTTGGCCAGCTCATCAGCTAGCCAG	5340
Qy	5341	CTGTCTGCGCTTCTTGAAGGCAACATGATGCTACCGCTCATGCTCCCGGAGGAGTGC	5400
Db	5341	CTGTCTGCGCTTCTTGAAGGCAACATGATGCTACCGCTCATGCTCCCGGAGGAGTGC	5400
Qy	5401	CTCATCGAGCCCAACCTCTCTGTCGGCGGAGAGATGGCGGGGAAACATCCCGGCTGGAG	5460
Db	5401	CTCATCGAGCCCAACCTCTCTGTCGGCGGAGAGATGGCGGGGAAACATCCCGGCTGGAG	5460
Qy	5461	TCAGAAATAAGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCCAAGCGGAGGAGTGC	5520
Db	5461	TCAGAAATAAGTAGTAAATTTTGGACTCTTTTCGAGCGCTCCCAAGCGGAGGAGTGC	5520
Qy	5521	AGGAGAGTATCGTTCGGCGGAGATCTCGGAGGTCCAGGAAATTTCCCTCGAGCGATG	5580
Db	5521	AGGAGAGTATCGTTCGGCGGAGATCTCGGAGGTCCAGGAAATTTCCCTCGAGCGATG	5580
Qy	5581	CCCATATGGCAGCGCCGGATTAACACCTTCCACTGTTAGAGTCTTGGAGGACCCGGAC	5640
Db	5581	CCCATATGGCAGCGCCGGATTAACACCTTCCACTGTTAGAGTCTTGGAGGACCCGGAC	5640

Qy	5641	TACGTCCTCCAGTGGGTACACGGGTGTCCATTTCCGCGCTGCCAAGCGCCCTCCGATACCA	5700
Db	5641		
Qy	5701	CTTCCAACGAGGAAGAGACGGTTGTCTCTGTACGAATCTACCGTGTCTTCTGCTTGGCG	5760
Db	5701		
Qy	5761	GAGCTCGCACAAAGACCTTCGCGACGCTCCGAATCGTCGGCGCTCGACAGCGGCACGGCA	5820
Db	5761		
Qy	5821	ACGGCCTCTCTGACACAGCCCTCCGACGACGGCGACCGGGATCCGAAGTTCAGTCGTAC	5880
Db	5821		
Qy	5881	TCTTCANTGCCCCCTTTGAGGGGGAGCGGGGGATCCCGATCTCAGCGACGGGTCTTGG	5940
Db	5881		
Qy	5941	TCTACCGTAAAGCAGGAGGCTAGTGCAGGACGTCGTCTGTGTCGATGTCCTACACATGG	6000
Db	5941		
Qy	6001	ACAGGCGCCTGATCAGCCATCGCTGCGGAGGAACCAAGCTGCCATCAATGACATG	6060
Db	6001		
Qy	6061	AGCAATCTTTGCTCCGTACACCAACTTGGTCTATGTCTACACATCTCGCAGCGCAAG	6120
Db	6061		
Qy	6121	CTCGCGCAGAAAGAGTCACTTTGACAGACTCGAGTCTCTGACGACCACTACCGGGAC	6180
Db	6121		
Qy	6181	GTGCTCAAGGAGATGAAGGCGAAAGCGTCCACAGTTAAAGCTAAACTTCTATCCGTGGAG	6240
Db	6181		
Qy	6241	GAAGCCTGTAAGCTGAGCGCCCCACATTCGGCGAGATCTAAATTCGCTATGGGCAAG	6300
Db	6241		
Qy	6301	GAGTCCGGAACTTATCCAGCAAGCGCGTTAAACACATCCGCTCCGCTGTGGAAAGACTTG	6360
Db	6301		
Qy	6361	CTGGAAGACACTGAGACACCAATTTGACACCAACCATCATGGCAAAAATGAGGTTTCTGC	6420
Db	6361		
Qy	6421	GTCCAAACAGAGAAGGGGGCGCAAGCCAGCTCGCCTATCGTATTCCAGATTTGGGG	6480
Db	6421		
Qy	6481	GTTCTGTGTCGGAATAATGGCCCTTTAAGATGTGTCCTCCACCTCCCTCAGGCGGTG	6540
Db	6481		
Qy	6541	ATGGGCTCTTCATACGATTCCAATATCTTCTCTGGAAGCGGGTCTGCTGTGAT	6600
Db	6541		
Qy	6601	GCCTGGAAGCGAAGAAATGCCCTATGGGCTTCGATATGACACCCGCTGTTTGACTCA	6660
Db	6601		
Qy	6661	ACGGTCACTGAGATGACATCCGTGTTGAGGAGTCAATCTACCAATGTGTGACTTGGCC	6720
Db	6661		
Qy	6721	CCCGAAGCGACAGGCGCAATAGGTCGTCTACAGAGCGGCTTTTACATCGGGGGCCCCCTG	6780

[illegible]



Db	1501	AACCCCCACCTCGGAGCAGGTCCTCTCTGGGCCAAAGSCACGCTGATATAAGATACACCT	1560
Qy	1561	GCRAAGGCGGCACAAACCCAGTCGCCAGTTGTGAGTTGGATAGTTGTGAAAGAGTCAAA	1620
Db	1561	GCRAAGGCGGCACAAACCCAGTCGCCAGTTGTGAGTTGGATAGTTGTGAAAGAGTCAAA	1620
Qy	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTCAAGGATGCCAGAAAGGTACCCCATCTGT	1680
Db	1621	TGGCTCTCTCAAGCGTATTCAACAGGGGCTCAAGGATGCCAGAAAGGTACCCCATCTGT	1680
Qy	1681	ATGGGATCTGATCTTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAA	1740
Db	1681	ATGGGATCTGATCTTGGGGCCTCGGTGCACATGCTTTACATGTGTTTAGTCGAGGTTAAAA	1740
Qy	1741	AACCTCTAGGCCCCCGGAAACAAGGGAGCGTGGTTTTCTTTGAAAAACACGATATATACC	1800
Db	1741	AACCTCTAGGCCCCCGGAAACAAGGGAGCGTGGTTTTCTTTGAAAAACACGATATATACC	1800
Qy	1801	ATGGCGCTATTACGGGCTACTCCCAACAGACGAGGCCCTACTTGGCTGCATCATCACT	1860
Db	1801	ATGGCGCTATTACGGGCTACTCCCAACAGACGAGGCCCTACTTGGCTGCATCATCACT	1860
Qy	1861	AGCCTCAACGCGCGGACAGGAAACCAAGTCCAGGGGAGGTCCAAAGTGTCTCCACCGCA	1920
Db	1861	AGCCTCAACGCGCGGACAGGAAACCAAGTCCAGGGGAGGTCCAAAGTGTCTCCACCGCA	1920
Qy	1921	ACAAATCTTCTTGGCGACTCGTCAATGGGCTGTGTGGACTGTCTATCATGTGTC	1980
Db	1921	ACAAATCTTCTTGGCGACTCGTCAATGGGCTGTGTGGACTGTCTATCATGTGTC	1980
Qy	1981	GGCTCAAAAGACCTTTGCTGGCGCCAAAGGGGCCCAATCAACCAAAATGTACACCAATGTGGAC	2040
Db	1981	GGCTCAAAAGACCTTTGCTGGCGCCAAAGGGGCCCAATCAACCAAAATGTACACCAATGTGGAC	2040
Qy	2041	CAGGACCTCTGTCGGCTGGCAAGCGCCCGCGGGCGGCTTCCTTGACACCATGCACTGTC	2100
Db	2041	CAGGACCTCTGTCGGCTGGCAAGCGCCCGCGGGCGGCTTCCTTGACACCATGCACTGTC	2100
Qy	2101	GSCAGCTCGGACCTTTACTTGGTCAACAGGCGATCCCGATGTCATTCCGGTGGCGCGCGG	2160
Db	2101	GSCAGCTCGGACCTTTACTTGGTCAACAGGCGATCCCGATGTCATTCCGGTGGCGCGCGG	2160
Qy	2161	GGCGACAGCGGGGAGCCTACTCTGCCCGAGGCCCGTCTCCTACTTTGAAGGGCTCTTCG	2220
Db	2161	GGCGACAGCGGGGAGCCTACTCTGCCCGAGGCCCGTCTCCTACTTTGAAGGGCTCTTCG	2220
Qy	2221	GGCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGCATCTTTTGGGCTGGCGTGTGC	2280
Db	2221	GGCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGCATCTTTTGGGCTGGCGTGTGC	2280
Qy	2281	ACCCGAGGGTTGGGAAGGGCGGTGGACTTTGTACCCGTTCGAGTCTATGGAAAACCATATG	2340
Db	2281	ACCCGAGGGTTGGGAAGGGCGGTGGACTTTGTACCCGTTCGAGTCTATGGAAAACCATATG	2340
Qy	2341	CGGTCCCGGCTTTCACGGGAACTCTGTCGCCCTCGCGCGGTACCGCAGACATTCACAGTG	2400
Db	2341	CGGTCCCGGCTTTCACGGGAACTCTGTCGCCCTCGCGCGGTACCGCAGACATTCACAGTG	2400
Qy	2401	GGCCATCTACACGCCCTTACTGGTAGCGGCGCAGAGCACTAAGGTGCGGCTTCGATATGCA	2460
Db	2401	GGCCATCTACACGCCCTTACTGGTAGCGGCGCAGAGCACTAAGGTGCGGCTTCGATATGCA	2460
Qy	2461	GGCCAAAGGTATAAGTGTCTGTAAACCGGTCCGTGCGCGGCAACCTTAGGTTTCGGG	2520
Db	2461	GGCCAAAGGTATAAGTGTCTGTGTAAACCGGTCCGTGCGCGGCAACCTTAGGTTTCGGG	2520
Qy	2521	GGGTATATCTCTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGGTAAAGACCATC	2580
Db	2521	GGGTATATCTCTAAGGCACATGGTATCGACCTTAACATCAGAACCGGGGTAAAGACCATC	2580
Qy	2581	ACCAAGGGTGCCTCCCAATCAAGTACTCCACTATAGCAAGTTTCTTGCCGACGGTGGTTC	2640
Db	2581	ACCAAGGGTGCCTCCCAATCAAGTACTCCACTATAGCAAGTTTCTTGCCGACGGTGGTTC	2640

QY	2641	TCTGGGGCGCCCTATGACATCAATAATATGATGAGTGCCTCACTCACTGACTCGACCCT	2700
DB	2641	TCTGGGGCGCCCTATGACATCAATAATATGATGAGTGCCTCACTCACTGACTCGACCCT	2700
QY	2701	ATCTCGGGCATCGCACAGTCTGGACCAAGCGGAGACGGCTGGAGCGGCACTCGTCGTG	2760
DB	2701	ATCTCGGGCATCGCACAGTCTGGACCAAGCGGAGACGGCTGGAGCGGCACTCGTCGTG	2760
QY	2761	CTGCCACCCTACGCTCCGGGATCGGTCA CCGTGCACATCCAAACATCGAGGAGGTG	2820
DB	2761	CTGCCACCCTACGCTCCGGGATCGGTCA CCGTGCACATCCAAACATCGAGGAGGTG	2820
QY	2821	GCTCTGTGCGACACTCGAGAAATCCCTTTTATGGCAAAAGCCATCCCATCGAGACCATC	2880
DB	2821	GCTCTGTGCGACACTCGAGAAATCCCTTTTATGGCAAAAGCCATCCCATCGAGACCATC	2880
QY	2881	AAGGGGGGAGGCACCTCATTTTTCTGCCATTCAAAGAAGAAATGTGATGAGTCTCGCGG	2940
DB	2881	AAGGGGGGAGGCACCTCATTTTTCTGCCATTCAAAGAAGAAATGTGATGAGTCTCGCGG	2940
QY	2941	AAGTGTCTCGGCTCGGACTCAATGTGTGACATATTACGGGGCCCTTGATGTATCCGTC	3000
DB	2941	AAGTGTCTCGGCTCGGACTCAATGTGTGTAGCATATTTACCGGGCCCTTGATGTATCCGTC	3000
QY	3001	ATACCAACTAGCGGAGAGCTCATTTGCTAGCAACGAGACGCTCTAATGACGGGCTTTACC	3060
DB	3001	ATACCAACTAGCGGAGAGCTCATTTGCTAGCAACGAGACGCTCTAATGACGGGCTTTACC	3060
QY	3061	GGCGATTTGCACTCAGTGATCGACTGCAATACATGTGTCTACCCAGACGTCGATCTCAGC	3120
DB	3061	GGCGATTTGCACTCAGTGATCGACTGCAATACATGTGTCTACCCAGACGTCGATCTCAGC	3120
QY	3121	CTGACCCGACCTTACCAATGTGACAGACACGCTGCCACAAGACGCGGTTCACGCTCG	3180
DB	3121	CTGACCCGACCTTACCAATGTGACAGACACGCTGCCACAAGACGCGGTTCACGCTCG	3180
QY	3181	CAGCGGGGAGGACGAGCTGTGTAGGGGAGGATGGGCA TTTACAGGTTTGTACTCCAGGA	3240
DB	3181	CAGCGGGGAGGACGAGCTGTGTAGGGGAGGATGGGCA TTTACAGGTTTGTACTCCAGGA	3240
QY	3241	GAA CGGGCCCTCGGGATGTTGCA TTTCTCGTTCGTGCGAGTGCTATGACGGGCTGT	3300
DB	3241	GAA CGGGCCCTCGGGATGTTGCA TTTCTCGTTCGTGCGAGTGCTATGACGGGCTGT	3300
QY	3301	GCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGGTTGGGGCTTACCTAAACACA	3360
DB	3301	GCTTGGTACGAGCTCACGCCCGCGAGACCTCAGTTAGGTTGGGGCTTACCTAAACACA	3360
QY	3361	CCAGGGTTGCCCGTCTGCGCAGGCAATCTGGAGTTC TGGGAGAGCGTCTTTACAGGCTC	3420
DB	3361	CCAGGGTTGCCCGTCTGCGCAGGCAATCTGGAGTTC TGGGAGAGCGTCTTTACAGGCTC	3420
QY	3421	ACCCACATAGACGCCCATTTCTCTGCCAGACTAAGCAGGCAGGAGACAACTTCCCTAC	3480
DB	3421	ACCCACATAGACGCCCATTTCTCTGCCAGACTAAGCAGGCAGGAGACAACTTCCCTAC	3480
QY	3481	CTGTGTAGCATACCGAGCTACGGTGTGCGCCAGGGCTCAGGCTCCACTCCATCGTGGAC	3540
DB	3481	CTGTGTAGCATACCGAGCTACGGTGTGCGCCAGGGCTCAGGCTCCACTCCATCGTGGAC	3540
QY	3541	CAAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCACGGGCCAAAGCCCTCTG	3600
DB	3541	CAAAATGTGGAAGTGTCTCATACGGCTAAAGCCTACGCTGCACGGGCCAAAGCCCTCTG	3600
QY	3601	TATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACCACACACCCCAATAACCAATACATC	3660
DB	3601	TATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACCACACACCCCAATAACCAATACATC	3660
QY	3661	ATGGCAATGATGTGGCTGACCTGTGAGTGTCTACGAGCACTCGGTGCTGGTAGGCGGA	3720
DB	3661	ATGGCAATGATGTGGCTGACCTGTGAGTGTCTACGAGCACTCGGTGCTGGTAGGCGGA	3720

QY	3721	GTCTAGACGCTCTGGCCGCGTATTGCTGACAAACAGGACGCGTGGTCAATTTGGGCGAG	3780		Db	4801	CAAAACACCTGCCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCCATGAGG	4860
Db	3721	GTCTAGACGCTCTGGCCGCGTATTGCTGACAAACAGGACGCGTGGTCAATTTGGGCGAG	3780		QY	4861	ATCGTGGGCGCTAGACCTGTAGTAACAGCTGGCATGGAACATTTCCCATTAACCGGTAC	4920
QY	3781	ATCATCTTGTCCGAAAGCCGCGCATCATTCGCGACAGGAAAGTCTTTACCGGAGTTTC	3840		Db	4861	ATCGTGGGCGCTAGACCTGTAGTAACAGCTGGCATGGAACATTTCCCATTAACCGGTAC	4920
Db	3781	ATCATCTTGTCCGAAAGCCGCGCATCATTCGCGACAGGAAAGTCTTTACCGGAGTTTC	3840		QY	4921	ACCAGGGCGCTGTCAGCGCTCTCCCGGCGCAAAATTTCTTAGGGCGCTGTGGCGGGTG	4980
QY	3841	GATGAGTGGAAAGTGGCCCTCACACCTCCCTTACATCGAACAGGAATGAGCTCGCC	3900		Db	4921	ACCAGGGCGCTGTCAGCGCTCTCCCGGCGCAAAATTTCTTAGGGCGCTGTGGCGGGTG	4980
Db	3841	GATGAGTGGAAAGTGGCCCTCACACCTCCCTTACATCGAACAGGAATGAGCTCGCC	3900		QY	4981	GCTGTGAGGAGTACGTGGAGTTACCGGGTGGGGGATTTCCACTAGTACGGGCGATG	5040
QY	3901	GAAATTTCAACACAGAGGCAATCGGTTGTGCAAAACAGCCACCAAGCAAGCGGAGGT	3960		Db	4981	GCTGTGAGGAGTACGTGGAGTTACCGGGTGGGGGATTTCCACTAGTACGGGCGATG	5040
Db	3901	GAAATTTCAACACAGAGGCAATCGGTTGTGCAAAACAGCCACCAAGCAAGCGGAGGT	3960		QY	5041	ACCACTGACAAAGTAAAGTGGCCCTGTGTCAGGTTCCGGCCCCCGGAATTTCTTACACAAAGTG	5100
QY	3961	GCTGTCCCGTGGTGGAAATCCAAGTGGCGGACCTCGAAGCCCTTCTGGGCGAAGCATATG	4020		Db	5041	ACCACTGACAAAGTAAAGTGGCCCTGTGTCAGGTTCCGGCCCCCGGAATTTCTTACACAAAGTG	5100
Db	3961	GCTGTCCCGTGGTGGAAATCCAAGTGGCGGACCTCGAAGCCCTTCTGGGCGAAGCATATG	4020		QY	5101	GATGGGGTCCGGTTCGACAGGTACGCTCCAGGTGCAAAACCCCTCTTACGGGAGGAGTTC	5160
QY	4021	TGGAATTTCAATCAGCGGATACAAATATTTAGCAGCTTTGTCACCTCTGCTGGACACCC	4080		Db	5101	GATGGGGTCCGGTTCGACAGGTACGCTCCAGGTGCAAAACCCCTCTTACGGGAGGAGTTC	5160
Db	4021	TGGAATTTCAATCAGCGGATACAAATATTTAGCAGCTTTGTCACCTCTGCTGGACACCC	4080		QY	5161	ACATTCCTGTGTCGCGCTCAATCAATACCTGTGGTCAACAGCTCCCATGCGAGCCGAA	5220
QY	4081	CGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCCCGCTCACACCCCAACAT	4140		Db	5161	ACATTCCTGTGTCGCGCTCAATCAATACCTGTGGTCAACAGCTCCCATGCGAGCCGAA	5220
Db	4081	CGATAGCATCACTGATGGCATTCACAGCTCTATCACAGCCCGCTCACACCCCAACAT	4140		QY	5221	CCGACGTAGAGTGTCTCACCTTCCATGTCTACACGACCCCTCCACATTCGCGCGAGAGC	5280
QY	4141	ACCTCTCTGTTTAAACATCTCTGGGGGATGGTGGCGCCCAACTTGTCTCCAGCGCT	4200		Db	5221	CCGACGTAGAGTGTCTCACCTTCCATGTCTACACGACCCCTCCACATTCGCGCGAGAGC	5280
Db	4141	ACCTCTCTGTTTAAACATCTCTGGGGGATGGTGGCGCCCAACTTGTCTCCAGCGCT	4200		QY	5281	GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTTGGCCAGTCAATAGCTAGCCAG	5340
QY	4201	GCTTCTGCTTTCTGATAGCGCGCATCGCTCGAGCGGCTGTGGCAGCATAGGCTTTGGG	4260		Db	5281	GCTAAGCGTAGGCTGGCGAGGGATCTCCCGCTCTTTGGCCAGTCAATAGCTAGCCAG	5340
Db	4201	GCTTCTGCTTTCTGATAGCGCGCATCGCTCGAGCGGCTGTGGCAGCATAGGCTTTGGG	4260		QY	5341	CTGTCTCGCTTCTTTGAAAGGCAACATGCACTACCCGTCTTCCCGGACGCTACAGCTGAC	5400
QY	4261	AAGTGCTTGTGATATTTTGGCAGTTTATGAGCAGGGGTGGCAGGCGCTCGTGGCC	4320		Db	5341	CTGTCTCGCTTCTTTGAAAGGCAACATGCACTACCCGTCTTCCCGGACGCTACAGCTGAC	5400
Db	4261	AAGTGCTTGTGATATTTTGGCAGTTTATGAGCAGGGGTGGCAGGCGCTCGTGGCC	4320		QY	5401	CTCATCGAGGCCAACCTCTCTGTGGCGCAGAGATGGCGGGAACATCACCCCGGTGGAG	5460
QY	4321	TTTAAGGTATGAGCGGAGATGCTCTCCACGAGGACCTGGTTAACTACTCCCTGCT	4380		Db	5401	CTCATCGAGGCCAACCTCTCTGTGGCGCAGAGATGGCGGGAACATCACCCCGGTGGAG	5460
Db	4321	TTTAAGGTATGAGCGGAGATGCTCTCCACGAGGACCTGGTTAACTACTCCCTGCT	4380		QY	5461	TCGAAATAAGGTAGTAAATTTTGGACTTTTCGAGCGCTTCCAAAGCGGAGGAGATGAG	5520
QY	4381	ATCTCTCTCCCTGCGCCCTAGTCTGGGGTCTGTGGCAGCGATCTGCTGCGCAC	4440		Db	5461	TCGAAATAAGGTAGTAAATTTTGGACTTTTCGAGCGCTTCCAAAGCGGAGGAGATGAG	5520
Db	4381	ATCTCTCTCCCTGCGCCCTAGTCTGGGGTCTGTGGCAGCGATCTGCTGCGCAC	4440		QY	5521	AGGAACTATCCGTTCCGCGGAGATCTCGGAGGTCCAGGAAATTTCCCTTCGAGCGATG	5580
QY	4441	GTGGCCCCAGGGAGGGGGCTGTGAGTGAATGAACCGGCTGTAGGGTTTCGCTTCGCGG	4500		Db	5521	AGGAACTATCCGTTCCGCGGAGATCTCGGAGGTCCAGGAAATTTCCCTTCGAGCGATG	5580
Db	4441	GTGGCCCCAGGGAGGGGGCTGTGAGTGAATGAACCGGCTGTAGGGTTTCGCTTCGCGG	4500		QY	5581	CCCATATGGGACGCGCCGATTAACACCTCTCACTGTTAGAGTCTCGAAGGAGCCCGGAC	5640
QY	4501	GSTAAACACGCTCTCCCGACGACTATGTGCTGAGAGCGACGTGAGCAGCGTCACT	4560		Db	5581	CCCATATGGGACGCGCCGATTAACACCTCTCACTGTTAGAGTCTCGAAGGAGCCCGGAC	5640
Db	4501	GSTAAACACGCTCTCCCGACGACTATGTGCTGAGAGCGACGTGAGCAGCGTCACT	4560		QY	5641	TAGCTCCCTCAAGTGTACACGGGTGTCCATTCGCGCTTCCCAAGGCGCCCTCCGATACCA	5700
QY	4561	CAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGGCTTCCACAGTGAATCAAC	4620		Db	5641	TAGCTCCCTCAAGTGTACACGGGTGTCCATTCGCGCTTCCCAAGGCGCCCTCCGATACCA	5700
Db	4561	CAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGGCTTCCACAGTGAATCAAC	4620		QY	5701	CTTCCAGGAGGAGGAGGAGGTTGCTGTCTGATCTAGATCTACCGTGTCTTCTGCTGGCG	5760
QY	4621	GAGACTGTCTCCAGCCATGCTCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC	4680		Db	5701	CTTCCAGGAGGAGGAGGAGGTTGCTGTCTGATCTAGATCTACCGTGTCTTCTGCTGGCG	5760
Db	4621	GAGACTGTCTCCAGCCATGCTCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC	4680		QY	5761	GAGCTGCCACAAAGACCTTTCGGCAGCTCCGAAATCTGCGCGCTCGACGCGGACGGCA	5820
QY	4681	ACGGTGTGACTGATTTCAAGACTGTGCTCAAGTCCAGTCTCTGCGCGGATTTGCCGGGA	4740		Db	5761	GAGCTGCCACAAAGACCTTTCGGCAGCTCCGAAATCTGCGCGCTCGACGCGGACGGCA	5820
Db	4681	ACGGTGTGACTGATTTCAAGACTGTGCTCAAGTCCAGTCTCTGCGCGGATTTGCCGGGA	4740		QY	5821	ACGGCTCTCTGACACGAGCCCTCCGACGCGGCGACGCGGATCCGACGTTGAGTCTGAT	5880
QY	4741	GTCCCTCTCTTCTCATGTCACGTTGGGTACAGGAGTCTGGCGGGCGACGATCATG	4800		Db	5821	ACGGCTCTCTTCTGACACGAGCCCTCCGACGCGGCGACGCGGATCCGACGTTGAGTCTGAT	5880
Db	4741	GTCCCTCTCTTCTCATGTCACGTTGGGTACAGGAGTCTGGCGGGCGACGATCATG	4800		QY	5881	TCTCTCATGCCCCCTTTCAGGGGAGCGCGGGATTCGCGATCTCAGCGACGCGGTCTTGG	5940
QY	4801	CAAAACACCTGCCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCCATGAGG	4860					





```
; Publication No. US20020155133A1
; GENERAL INFORMATION:
; APPLICANT: ANADYS Pharmaceuticals, Inc.
; APPLICANT: Bichko, Vadim
; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPLICATION
; FILE REFERENCE: 0342/1H3950S1
; CURRENT APPLICATION NUMBER: US/10/005,469
; CURRENT FILING DATE: 2002-04-18
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV Replicon from cell line HCVr24
US-10-005-469-6

Query Match          100.0%; Score 7985.8; DB 13; Length 7992;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 7987; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GCCAGCCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGAACTACTG 60
DB 1 GCCAGCCCCCGATTGGGGCGACACTCCACCATAGATCACTCCCTGTGAGAACTACTG 60

QY 61 TCITCAGCAGAAAGCGTCTAGCCATGCGCGTTAGTATGAGTGTGCGTGAGCCTCCAGGAC 120
DB 61 TCITCAGCAGAAAGCGTCTAGCCATGCGCGTTAGTATGAGTGTGCGTGAGCCTCCAGGAC 120

QY 121 CCCCCCTCCCGGAGAGCCTAGTGGTCTCGGAGACCGGTGAGTACACCGAAATGGCAG 180
DB 121 CCCCCCTCCCGGAGAGCCTAGTGGTCTCGGAGACCGGTGAGTACACCGAAATGGCAG 180

QY 181 GACGACCGGGTCTTTCTTTGATCAACCCCGTCAATGCTCGAGATTGGGCGTGCCTCC 240
DB 181 GACGACCGGGTCTTTCTTTGATCAACCCCGTCAATGCTCGAGATTGGGCGTGCCTCC 240

QY 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGCCCTTGTGACTGCGCTGATAGG 300
DB 241 GCGAGACTGTAGCCGAGTAGTGTGGTTCGGAAGCCCTTGTGACTGCGCTGATAGG 300

QY 301 GTGCTTCGAGTGCCTCCGAGGCTCTGTAGACCGGTGACCATGACGACGAACTCTAAAC 360
DB 301 GTGCTTCGAGTGCCTCCGAGGCTCTGTAGACCGGTGACCATGACGACGAACTCTAAAC 360

QY 361 CTCAAGAAACAAACAAAGGCGCGCCATGATTGAACAAAGATGGATTGACGAGGTTCTC 420
DB 361 CTCAAGAAACAAACAAAGGCGCGCCATGATTGAACAAAGATGGATTGACGAGGTTCTC 420

QY 421 CGGCGCTTGGTGGAGAGGCTATTGCGCTATGCTGGGCAACACAGCAATCGGCTGCT 480
DB 421 CGGCGCTTGGTGGAGAGGCTATTGCGCTATGCTGGGCAACACAGCAATCGGCTGCT 480

QY 481 CTGATGCGCGCTGTTCCGCGTGTACGCGAGGCGCGCGCTTCTTTTGTCAAGACCG 540
DB 481 CTGATGCGCGCTGTTCCGCGTGTACGCGAGGCGCGCGCTTCTTTTGTCAAGACCG 540

QY 541 ACCTGTCCGGTGCCTGAATGAACCTGACGAGCAGGCGCGCGCTATCGTGGCTGSCCA 600
DB 541 ACCTGTCCGGTGCCTGAATGAACCTGACGAGCAGGCGCGCGCTATCGTGGCTGSCCA 600

QY 601 CGAGGCGCTTCTTGGCGAGTGTGTCTGACGTTGTCACTGAAGCGGGGAAGGACTGGC 660
DB 601 CGAGGCGCTTCTTGGCGAGTGTGTCTGACGTTGTCACTGAAGCGGGGAAGGACTGGC 660

QY 661 TGTCTATTGGCGAAGTCCCGGCGAGATCTCTGTCTCATCTCACTTGTCTTCCGCGAGA 720
DB 661 TGTCTATTGGCGAAGTCCCGGCGAGATCTCTGTCTCATCTCACTTGTCTTCCGCGAGA 720

QY 721 AAGTATCCATCATGGCTGATGCAATGGCGGCTGCATACGCTTGATCCGGCTACCTGCC 780
```

```
DB 721 AAGTATCCATCATGGCTGATGCAATGGCGGCTGCATACGCTTGTATCCGGTACCTGCC 780
QY 781 CATTCGACCAACAAAGCAAAATCCCATCGAGCGAGCAGTACTCGATGGAAGCGGTC 840
DB 781 CATTCGACCAACAAAGCAAAATCCCATCGAGCGAGCAGTACTCGATGGAAGCGGTC 840
QY 841 TTGTGATCAGGATGATCTGACGAAAGAGCATCAGGGGCTCGCGCAGCGAATGTTGC 900
DB 841 TTGTGATCAGGATGATCTGACGAAAGAGCATCAGGGGCTCGCGCAGCGAATGTTGC 900
QY 901 CCAGGCTCAAGCGCGCATGCCGAGCGGAGGATCTCTGTGATGCCATGCGGATGCT 960
DB 901 CCAGGCTCAAGCGCGCATGCCGAGCGGAGGATCTCTGTGATGCCATGCGGATGCT 960
QY 961 GCTTGCAGATATCATGGTGAATAATGCCGCTTTTCTGGATTCTATCGACTGTGCCGCG 1020
DB 961 GCTTGCAGATATCATGGTGAATAATGCCGCTTTTCTGGATTCTATCGACTGTGCCGCG 1020
QY 1021 TGGGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATGCTGAAGAGC 1080
DB 1021 TGGGTGCGGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATGCTGAAGAGC 1080
QY 1081 TTGGCGGGAATGGGCTGACCGCTTCTCTGCTTACGGTATCGCCGCTCCGATTCG 1140
DB 1081 TTGGCGGGAATGGGCTGACCGCTTCTCTGCTTACGGTATCGCCGCTCCGATTCG 1140
QY 1141 AGCGATCGCCTTCTATCGCCTTCTTACGAGTCTTCTGAGTTTAAACAGACCAACAG 1200
DB 1141 AGCGATCGCCTTCTATCGCCTTCTTACGAGTCTTCTGAGTTTAAACAGACCAACAG 1200
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCCCCCCCTTAAGCTTACTGCC 1260
DB 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTCCCTCCCCCCCCCTTAAGCTTACTGCC 1260
QY 1261 CGAAGCGCTTGAATAGCGGCTGCTGCTTCTATATGTTTATTTTCCACCATATG 1320
DB 1261 CGAAGCGCTTGAATAGCGGCTGCTGCTTCTATATGTTTATTTTCCACCATATG 1320
QY 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCTCTGTTCTTGTAGAGCATTCCT 1380
DB 1321 CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTGCGCTCTGTTCTTGTAGAGCATTCCT 1380
QY 1381 AGGCTCTTTCCCTCTCGCAAAGGAATGCAAGGTCTGTGATGTCGTAAGGAAGCA 1440
DB 1381 AGGCTCTTTCCCTCTCGCAAAGGAATGCAAGGTCTGTGATGTCGTAAGGAAGCA 1440
QY 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACAGTCTGTAGGACCTTTTGAGGACGCG 1500
DB 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACAGTCTGTAGGACCTTTTGAGGACGCG 1500
QY 1501 AACCCCGCCTCGGACAGGCTCTCTGCGGCAAAAGCCAGCTGTATAGATACACCT 1560
DB 1501 AACCCCGCCTCGGACAGGCTCTCTGCGGCAAAAGCCAGCTGTATAGATACACCT 1560
QY 1561 GAAAGCGGCAACACCCAGTGCCTGTGAGTTGGATAGTGTGGAAGAGTCAAA 1620
DB 1561 GAAAGCGGCAACACCCAGTGCCTGTGAGTTGGATAGTGTGGAAGAGTCAAA 1620
QY 1621 TGGCTCTCTCAAGCTATTCACCAAGGCGCTGAAGATGCCAGAGGTACCCCATTTG 1680
DB 1621 TGGCTCTCTCAAGCTATTCACCAAGGCGCTGAAGATGCCAGAGGTACCCCATTTG 1680
QY 1681 ATGGATCTGATCTGGGCTCGTGACATGTTTACATGCTTTAGTCGAGGTTAAAA 1740
DB 1681 ATGGATCTGATCTGGGCTCGTGACATGTTTACATGCTTTAGTCGAGGTTAAAA 1740
QY 1741 AAGCTCTAGGCGCCCGAACCAAGGCGAGTGGTTTCTTTTGAAGAAACCGATAATACC 1800
DB 1741 AAGCTCTAGGCGCCCGAACCAAGGCGAGTGGTTTCTTTTGAAGAAACCGATAATACC 1800
QY 1801 ATGGCGCTTATACGCGCTTCTCCCAACAGACGCGAGGCTTCTTGGCTGATCATCACT 1860
```

Db 1801 ATGGCGCTATTACGGCTACTCCCAACAGACGCGAGGCGCTACTTTGGTGCATCATCACT 1860  
Qy 1861 AGCTTCAAGGCGGGGACAGGAACCAAGTCTGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
Db 1861 AGCTTCAAGGCGGGGACAGGAACCAAGTCTGAGGGGAGGTCCAAGTGTCTCCACCGCA 1920  
Qy 1921 ACACAACTTCTCTGGGACCTCGCTCAATGCGGTGTGTTGACGTGTCTATCATGTGTC 1980  
Db 1921 ACACAACTTCTCTGGGACCTCGCTCAATGCGGTGTGTTGACGTGTCTATCATGTGTC 1980  
Qy 1981 GGCTCAAGACCTTGTCCGGCCCAAGGCGCCAAATCACCCAAATGTACCAATGTGGAC 2040  
Db 1981 GGCTCAAGACCTTGTCCGGCCCAAGGCGCCAAATCACCCAAATGTACCAATGTGGAC 2040  
Qy 2041 CAGGACCTCGTGTGGCAAGGCGCCCGGGGGCGGTCTTGTGACACATGTGACCTGG 2100  
Db 2041 CAGGACCTCGTGTGGCAAGGCGCCCGGGGGCGGTCTTGTGACACATGTGACCTGG 2100  
Qy 2101 GGCAGCTCGGACCTTTACTTGTTCAGAGGATGCGATGTCAATCCGGTGGCGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTGTTCAGAGGATGCGATGTCAATCCGGTGGCGCGCGG 2160  
Qy 2161 GGCAGCAGAGGGGAGCTACTCTCCCGAGCGCGTCTCTCTTGAAGGCTCTTCG 2220  
Db 2161 GGCAGCAGAGGGGAGCTACTCTCCCGAGCGCGTCTCTCTTGAAGGCTCTTCG 2220  
Qy 2221 GGCAGTCCACTGTCTGTGCGCTCGGGGACGCTGTGGGATCTTTGGGTCGCGTGTG 2280  
Db 2221 GGCAGTCCACTGTCTGTGCGCTCGGGGACGCTGTGGGATCTTTGGGTCGCGTGTG 2280  
Qy 2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGGTGAGTCTATGGAACCACTATG 2340  
Db 2281 ACCGAGGGGTTGCGAAGCGGTGGACTTTGTACCGGTGAGTCTATGGAACCACTATG 2340  
Qy 2341 CGGTCCCGGTCTTACGGAACAACCTCGTCCCTCCGCGCGTACCGAGACATTCAGGTG 2400  
Db 2341 CGGTCCCGGTCTTACGGAACAACCTCGTCCCTCCGCGCGTACCGAGACATTCAGGTG 2400  
Qy 2401 GCCATCTACGCGCCCTACTGGTAGGGCAAGACGACTAAGGTGCGGTGGTATGCA 2460  
Db 2401 GCCATCTACGCGCCCTACTGGTAGGGCAAGACGACTAAGGTGCGGTGGTATGCA 2460  
Qy 2461 GCCAAGGTTAAGGTGCTTCTCAACCGGTCCGTCCGCGCACTAGGTTCGGG 2520  
Db 2461 GCCAAGGTTAAGGTGCTTCTCAACCGGTCCGTCCGCGCACTAGGTTCGGG 2520  
Qy 2521 GCCTATATGTCTAAGGCACATGTPATCGACCTTAACATCAGAACCGGGGTAAAGCACATC 2580  
Db 2521 GCCTATATGTCTAAGGCACATGTPATCGACCTTAACATCAGAACCGGGGTAAAGCACATC 2580  
Qy 2581 ACCACGGTGGCCCATCAGTACTCCACTATGGCAAGTTCCTTGGCGAGGTGTTGC 2640  
Db 2581 ACCACGGTGGCCCATCAGTACTCCACTATGGCAAGTTCCTTGGCGAGGTGTTGC 2640  
Qy 2641 TCTGGGGCGCTATGATCATATAATATGTAGTGCCACTCAACTGACTCGACCACT 2700  
Db 2641 TCTGGGGCGCTATGATCATATAATATGTAGTGCCACTCAACTGACTCGACCACT 2700  
Qy 2701 ATCTGGGCGATCGGCACAGTCTCGGACCAAGCGGAGACGCTGAGCGGAGCTCGTGTG 2760  
Db 2701 ATCTGGGCGATCGGCACAGTCTCGGACCAAGCGGAGACGCTGAGCGGAGCTCGTGTG 2760  
Qy 2761 CTGCGCACCGTACGCTCCGGGATCGGTACCGTGCACATCCAAACATCGAGGAGTG 2820  
Db 2761 CTGCGCACCGTACGCTCCGGGATCGGTACCGTGCACATCCAAACATCGAGGAGTG 2820  
Qy 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCCATCGAGACCATC 2880  
Db 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCCATCCCCATCGAGACCATC 2880  
Qy 2881 AAGGGGGGAGGCACTCATTTTCTGCCATTCAGAGAAATGTGATGAGCTCGCGCG 2940  
Db 2881 AAGGGGGGAGGCACTCATTTTCTGCCATTCAGAGAAATGTGATGAGCTCGCGCG 2940

Qy 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTCAATGATCCGTC 3000  
Db 2941 AAGCTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTCAATGATCCGTC 3000  
Qy 3001 ATACCAACTAGCGGAGACGTCTATTGTCTAGCAACGAGCGCTCTAATGACGGCTTACC 3060  
Db 3001 ATACCAACTAGCGGAGACGTCTATTGTCTAGCAACGAGCGCTCTAATGACGGCTTACC 3060  
Qy 3061 GCGGATTTGCACTCAGTGTATGACTGCAATATATGTGTCAACAGACAGTCCGACTTCAG 3120  
Db 3061 GCGGATTTGCACTCAGTGTATGACTGCAATATATGTGTCAACAGACAGTCCGACTTCAG 3120  
Qy 3121 CTGGACCCGACCTTCAACCAATTGAGACGACGCGTCCCAAGACGCGGTGTACGCTCG 3180  
Db 3121 CTGGACCCGACCTTCAACCAATTGAGACGACGCGTCCCAAGACGCGGTGTACGCTCG 3180  
Qy 3181 CAGCGGCGAGGACGACGCTGTGTAGGGGAGGATGGGCAATTACAGTTTGTGACTCCAGA 3240  
Db 3181 CAGCGGCGAGGACGACGCTGTGTAGGGGAGGATGGGCAATTACAGTTTGTGACTCCAGA 3240  
Qy 3241 GAACGSCCTCGGGCATGTTCAATTCTCGGTTCTGTGCGAGTGTCTATGACGCGGCTGT 3300  
Db 3241 GAACGSCCTCGGGCATGTTCAATTCTCGGTTCTGTGCGAGTGTCTATGACGCGGCTGT 3300  
Qy 3301 GCTTGTGTAGAGCTCACGCGCGCGGAGACCTCAGTTTAGTTGCGGCTTACCTTAAACACA 3360  
Db 3301 GCTTGTGTAGAGCTCACGCGCGCGGAGACCTCAGTTTAGTTGCGGCTTACCTTAAACACA 3360  
Qy 3361 CAGGGTTGCCGTCTGCGACGACCATCTGGAGTCTTGGGAGAGCGCTTACAGGCTC 3420  
Db 3361 CAGGGTTGCCGTCTGCGACGACCATCTGGAGTCTTGGGAGAGCGCTTACAGGCTC 3420  
Qy 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTAAGCAGGAGAGACAACTTCCCTAC 3480  
Db 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTAAGCAGGAGAGACAACTTCCCTAC 3480  
Qy 3481 CTGTGTAGTATACAGGCTACGCTGTGCGGAGGCTCAGGCTCCACTCCTCATCTGTGGAC 3540  
Db 3481 CTGTGTAGTATACAGGCTACGCTGTGCGGAGGCTCAGGCTCCACTCCTCATCTGTGGAC 3540  
Qy 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACGCGGCAACGCGCTGCTG 3600  
Db 3541 CAAATGTGGAAGTGTCTCATACGGCTAAAGCTTACGCTGACGCGGCAACGCGCTGCTG 3600  
Qy 3601 TATAGGCTGGAGCGGTTCAAAACGAGGTTACTACACACACACCCCAATAAATAATC 3660  
Db 3601 TATAGGCTGGAGCGGTTCAAAACGAGGTTACTACACACACACCCCAATAAATAATC 3660  
Qy 3661 ATGGCATGTATGTCCGTGACCTGGAGGTCGTCAAGACACCTGGGTGTGTAGGCGGA 3720  
Db 3661 ATGGCATGTATGTCCGTGACCTGGAGGTCGTCAAGACACCTGGGTGTGTAGGCGGA 3720  
Qy 3721 GTCTTAGCAGCTGTGGCGGTATTGCTGACACAGCAGCGGTGTCATTGTGGGAGG 3780  
Db 3721 GTCTTAGCAGCTGTGGCGGTATTGCTGACACAGCAGCGGTGTCATTGTGGGAGG 3780  
Qy 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTTCCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Qy 3841 GATGAGATGGAAGAGTGCCTCACCTCCCTTACATCGAAACAGGGAATGACGCTGCC 3900  
Db 3841 GATGAGATGGAAGAGTGCCTCACCTCCCTTACATCGAAACAGGGAATGACGCTGCC 3900  
Qy 3901 GAACAAATCAACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGCT 3960  
Db 3901 GAACAAATCAACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGCT 3960  
Qy 3961 GCTCTCCGTTGGTGAATCCAGTGGCGGACCCCTCGAAGCTTCTTGGCGGAAGCATATG 4020  
Db 3961 GCTCTCCGTTGGTGAATCCAGTGGCGGACCCCTCGAAGCTTCTTGGCGGAAGCATATG 4020

QY 4021 TGGAAATTCATCAGCGGGATACAATAATTTAGAGGCTTGTGCACCTCTGCTGGCAACCC 4080  
Db 4021 TGGAAATTCATCAGCGGGATACAATAATTTAGAGGCTTGTGCACCTCTGCTGGCAACCC 4080  
QY 4081 GGGATAGCATCACTGATGCGCATTCACAGCTCTATCACCAGCGCGTCAACCCAAACAT 4140  
Db 4081 GGGATAGCATCACTGATGCGCATTCACAGCTCTATCACCAGCGCGTCAACCCAAACAT 4140  
QY 4141 ACCCTCCTGTTTAAACATCTCGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
Db 4141 ACCCTCCTGTTTAAACATCTCGGGGGATGGGTGGCGCCCAACTTGTCTCTCCAGCGCT 4200  
QY 4201 GCTTCTGCTTTCGTAGGCGCGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCCCTTGGG 4260  
Db 4201 GCTTCTGCTTTCGTAGGCGCGCATCGCTGGAGCGGCTGTTGGCAGCATAGGCCCTTGGG 4260  
QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCGCTGCTGGCC 4320  
Db 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGCGCGCTGCTGGCC 4320  
QY 4321 TTTAAGGTCAATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
Db 4321 TTTAAGGTCAATGAGCGCGAGATGCCCTCCACCGAGGACCTGGTTAACTCTCCCTGCT 4380  
QY 4381 ATCCCTCTCCCTGGCGCCCTAGTCGTCGGGTCTGTGGCAGCATACTGGTCGGCAC 4440  
Db 4381 ATCCCTCTCCCTGGCGCCCTAGTCGTCGGGTCTGTGGCAGCATACTGGTCGGCAC 4440  
QY 4441 GTGGGCCAGGGAGGGGGCTGTGCAATGATGAACCGGCTGATAGCGTTTCGTCGGG 4500  
Db 4441 GTGGGCCAGGGAGGGGGCTGTGCAATGATGAACCGGCTGATAGCGTTTCGTCGGG 4500  
QY 4501 GGTAAACACAGCTCTCCCGCAGCACTATGTGCTGAGCGGAGCTGACGACGTGTCACT 4560  
Db 4501 GGTAAACACAGCTCTCCCGCAGCACTATGTGCTGAGCGGAGCTGACGACGTGTCACT 4560  
QY 4561 CAGATCCCTCTAGTCTTACATCACTCAGCTGTGTAAGAGGCTTCAACAGTGAATCAAC 4620  
Db 4561 CAGATCCCTCTAGTCTTACATCACTCAGCTGTGTAAGAGGCTTCAACAGTGAATCAAC 4620  
QY 4621 GAGGACTGCTCCAGCGCATGCTCGGCTGCTGGTCTGAGATGATGATGATG 4680  
Db 4621 GAGGACTGCTCCAGCGCATGCTCGGCTGCTGGTCTGAGATGATGATGATGATG 4680  
QY 4681 ACGGTGTTGACTGATTTCAAGACCTGCTCAGTCCAGCTCCAGCTCCAGCTCCAGCT 4740  
Db 4681 ACGGTGTTGACTGATTTCAAGACCTGCTCAGTCCAGCTCCAGCTCCAGCTCCAGCT 4740  
QY 4741 GTCCCTCTCTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGCGAGCGCATCATG 4800  
Db 4741 GTCCCTCTCTCTCATGTCAACGTGGGTACAAGGAGTCTGGCGGGCGAGCGCATCATG 4800  
QY 4801 CAACACACCTGCCCATGTGAGCAGACATCAACCGGACATGTGAAAAACGGTTCCATGAGG 4860  
Db 4801 CAACACACCTGCCCATGTGAGCAGACATCAACCGGACATGTGAAAAACGGTTCCATGAGG 4860  
QY 4861 ATCGTGGGCTTAGGACCTGTAGTAAACCTGCAATGCAATTCACCATTAACGGTAC 4920  
Db 4861 ATCGTGGGCTTAGGACCTGTAGTAAACCTGCAATGCAATTCACCATTAACGGTAC 4920  
QY 4921 ACCACGGGCCCCGACCGCTCCCGGGCGCCAAATTAATCTAGGGCGCTGTGGCGGGTG 4980  
Db 4921 ACCACGGGCCCCGACCGCTCCCGGGCGCCAAATTAATCTAGGGCGCTGTGGCGGGTG 4980  
QY 4981 GCTGCTGAGAGTACGTGGAGGTTACCGGGTGGGGATTTCACCTAGCTGAGCGGGCATG 5040  
Db 4981 GCTGCTGAGAGTACGTGGAGGTTACCGGGTGGGGATTTCACCTAGCTGAGCGGGCATG 5040  
QY 5041 ACCACTGACAACTGAAGTCCCGTGTCAAGTTCCGCCCCCGAAATTTCTTCAAGAGTG 5100  
Db 5041 ACCACTGACAACTGAAGTCCCGTGTCAAGTTCCGCCCCCGAAATTTCTTCAAGAGTG 5100  
QY 5101 GATGGGCTGGGTGTCACAGGTACGCTCCAGCGTGCACAAACCCCTCTACGGGAGGAGT 5160

Db 5101 GATGGGCTGGGTGTCACAGGTACGCTCCAGCGTGCACAAACCCCTCTACGGGAGGAGT 5160  
QY 5161 ACATTCCTGGTGGTGGTCAATCAATACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 5220  
Db 5161 ACATTCCTGGTGGTGGTCAATCAATACCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 5220  
QY 5221 CCGGAGTAGCATGCTCATCTTCATGCTCACCGACCCCTCCCACTTACGGCGGAGAG 5280  
Db 5221 CCGGAGTAGCATGCTCATCTTCATGCTCACCGACCCCTCCCACTTACGGCGGAGAG 5280  
QY 5281 GCTAAGCGTAGGCTGGCGAGGGGATCTCCCGCTCTGCTGGCGAGCTCATCTAGCTAGC 5340  
Db 5281 GCTAAGCGTAGGCTGGCGAGGGGATCTCCCGCTCTGCTGGCGAGCTCATCTAGCTAGC 5340  
QY 5341 CTGTCTGGCGCTTCTTGAAGGCAACATGCTACCTCCGCTCATGCTCCCGAGCGCTGAC 5400  
Db 5341 CTGTCTGGCGCTTCTTGAAGGCAACATGCTACCTCCGCTCATGCTCCCGAGCGCTGAC 5400  
QY 5401 CTCATCGAGGCAACCTCTCTGCGCGCAGGAGATGGCGGGAAATCACCCCGCTGGAG 5460  
Db 5401 CTCATCGAGGCAACCTCTCTGCGCGCAGGAGATGGCGGGAAATCACCCCGCTGGAG 5460  
QY 5461 TCAGAAATTAAGGTAGTAAATTTTGGACTCTTTCGAGCGCTCCCAAGCGGAGGATGAG 5520  
Db 5461 TCAGAAATTAAGGTAGTAAATTTTGGACTCTTTCGAGCGCTCCCAAGCGGAGGATGAG 5520  
QY 5521 AGGAAAGTATCCGTTCCGGCGGAGATCTCGGAGGTCACGAAATTCCTCGAGCGATG 5580  
Db 5521 AGGAAAGTATCCGTTCCGGCGGAGATCTCGGAGGTCACGAAATTCCTCGAGCGATG 5580  
QY 5581 CCCATATGGGCAACCCCGGATTAACAACCTCTTCTAGAGTCTTGGAAAGACCCGAC 5640  
Db 5581 CCCATATGGGCAACCCCGGATTAACAACCTCTTCTAGAGTCTTGGAAAGACCCGAC 5640  
QY 5641 TACGTCCTCCAGTGGTACAGGCTGCTTTCGCGCTGCTGCAAGGCGCTTCCGATACCA 5700  
Db 5641 TACGTCCTCCAGTGGTACAGGCTGCTTTCGCGCTGCTGCAAGGCGCTTCCGATACCA 5700  
QY 5701 CCTCCAGGAGGAGAGAGCGTTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5760  
Db 5701 CCTCCAGGAGGAGAGAGCGTTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5760  
QY 5761 GAGTCTGCCACAAAGACCTTTCGCGAGCTCGAATCGTGGCGCTGCGACGCGGACCGCA 5820  
Db 5761 GAGTCTGCCACAAAGACCTTTCGCGAGCTCGAATCGTGGCGCTGCGACGCGGACCGCA 5820  
QY 5821 ACGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5880  
Db 5821 ACGGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5880  
QY 5881 TCCTTCCATGCCCCCTTGAAGGAGGAGCGCGGATCCCGATCTCAGCGAGCGGCTGCTGG 5940  
Db 5881 TCCTTCCATGCCCCCTTGAAGGAGGAGCGCGGATCCCGATCTCAGCGAGCGGCTGCTGG 5940  
QY 5941 TCTACCGTAAAGCAGGAGGCTAGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6000  
Db 5941 TCTACCGTAAAGCAGGAGGCTAGTGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 6000  
QY 6001 ACAGGCGCTGATCAGCGCATGCGCTGGCGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
Db 6001 ACAGGCGCTGATCAGCGCATGCGCTGGCGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
QY 6061 AGCAACTCTTGTCTCGTCCGTCACCACTTGTCTATGCTACACATCTCGCAGCGCAAGC 6120  
Db 6061 AGCAACTCTTGTCTCGTCCGTCACCACTTGTCTATGCTACACATCTCGCAGCGCAAGC 6120  
QY 6121 CTGCGCAGAGAAGGTCACTTTTGAAGACTGCGGCTGCTGCGAGCAGCAGTACCGGGAC 6180  
Db 6121 CTGCGCAGAGAAGGTCACTTTTGAAGACTGCGGCTGCTGCGAGCAGCAGTACCGGGAC 6180  
QY 6181 GTGCTCAAGGAGATGAAGCGGAGGCTCCAAGTTAAGGCTAAATCTTATCTCGTGGAG 6240

Db 6181 GTGCTCAGGAGATGAAGCGGAGCGCTCCACAGTTTAGGCTAAACTTCTATCCGTGGAG 6240  
QY |||||  
Db 6241 GAAGCCTGTAAAGTGAACGCCGCCCAATTGCGGCCAGATCTAAATTTGGCTATGGGCCAAAG 6300  
QY |||||  
Db 6241 GAAGCCTGTAAAGTGAACGCCGCCCAATTGCGGCCAGATCTAAATTTGGCTATGGGCCAAAG 6300  
QY |||||  
Db 6301 GACGTCCGGAACCTATCCAGCAAGCGCTTAACCAATCCGCTCGGTGGAAGGACTTG 6360  
QY |||||  
Db 6301 GACGTCCGGAACCTATCCAGCAAGCGCTTAACCAATCCGCTCGGTGGAAGGACTTG 6360  
QY |||||  
Db 6361 CTGGAAGACACTGAGACACCAATTTGACACACCATCATGCGCAAAAATAGAGTTTCTGC 6420  
QY |||||  
Db 6361 CTGGAAGACACTGAGACACCAATTTGACACACCATCATGCGCAAAAATAGAGTTTCTGC 6420  
QY |||||  
Db 6421 GTTCAACCAAGAGAGCGGGGCCGCAAGCCAGTCTGCTTATCGTATTTCCAGATTTGGGG 6480  
QY |||||  
Db 6421 GTTCAACCAAGAGAGCGGGGCCGCAAGCCAGTCTGCTTATCGTATTTCCAGATTTGGGG 6480  
QY |||||  
Db 6481 GTTCGTGTGCGAGAAATGGCCCTTTACGATGTGCTCCACCCCTCCCTCAGGCCGTG 6540  
QY |||||  
Db 6481 GTTCGTGTGCGAGAAATGGCCCTTTACGATGTGCTCCACCCCTCCCTCAGGCCGTG 6540  
QY |||||  
Db 6541 ATGGGCTCTTACGATTTCAATCTCTCTGGACAGCGGTCGAGTTCTCTGTGAAT 6600  
QY |||||  
Db 6541 ATGGGCTCTTACGATTTCAATCTCTCTGGACAGCGGTCGAGTTCTCTGTGAAT 6600  
QY |||||  
Db 6601 GCCTGAAACGGAAGAAATGCCCTATGGGCTTCGCATATGACACCCGCTGTTTGTACTCA 6660  
QY |||||  
Db 6601 GCCTGAAACGGAAGAAATGCCCTATGGGCTTCGCATATGACACCCGCTGTTTGTACTCA 6660  
QY |||||  
Db 6661 ACGTCACTGAGATGACATCCGCTGTTGAGGATCAATCTACCAATGTTGTACTTGGCC 6720  
QY |||||  
Db 6661 ACGTCACTGAGATGACATCCGCTGTTGAGGATCAATCTACCAATGTTGTACTTGGCC 6720  
QY |||||  
Db 6721 CCGGAAGCAGACAGGCAATTAAGTCTGCTCACAGAGCGGCTTTACATCGGGGCCCCCTG 6780  
QY |||||  
Db 6721 CCGGAAGCAGACAGGCAATTAAGTCTGCTCACAGAGCGGCTTTACATCGGGGCCCCCTG 6780  
QY |||||  
Db 6781 ACTAATTTCTAAAGGCGAGAACTCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
QY |||||  
Db 6781 ACTAATTTCTAAAGGCGAGAACTCGGCTATCGCGGTGCGCGGAGCGGTGACTGACG 6840  
QY |||||  
Db 6841 ACCAGTGGGTAATACCTTCAATGTTACTTCAAGGCCGCTCGGCTGCTGAGCTGCG 6900  
QY |||||  
Db 6841 ACCAGTGGGTAATACCTTCAATGTTACTTCAAGGCCGCTCGGCTGCTGAGCTGCG 6900  
QY |||||  
Db 6901 AAGCTCCAGGACTGCACGATGCTGATGCGGAGACGACCTTGTCTGTTACTGTGAAGC 6960  
QY |||||  
Db 6901 AAGCTCCAGGACTGCACGATGCTGATGCGGAGACGACCTTGTCTGTTACTGTGAAGC 6960  
QY |||||  
Db 6961 GCGGGACCCAAAGAGGAGGAGCGCTACGGGCTTTACGGAGGCTATGACTAGATAC 7020  
QY |||||  
Db 6961 GCGGGACCCAAAGAGGAGGAGCGCTACGGGCTTTACGGAGGCTATGACTAGATAC 7020  
QY |||||  
Db 7021 TCTGCCCTTGGGACCGGCCCAACCAACCAATACGACTTGGAGTTGATAACATATGC 7080  
QY |||||  
Db 7021 TCTGCCCTTGGGACCGGCCCAACCAACCAATACGACTTGGAGTTGATAACATATGC 7080  
QY |||||  
Db 7081 TCCTCCCAATGTGTCACTGCGGACGATGCTGCTGCGCAAAAGGGTGTACTATCTCACCGT 7140  
QY |||||  
Db 7081 TCCTCCCAATGTGTCACTGCGGACGATGCTGCTGCGCAAAAGGGTGTACTATCTCACCGT 7140  
QY |||||  
Db 7141 GACCCACACCCCTTGGCGGGCTGCGTGGAGACAGTACACACACTCCAGTCAAT 7200  
QY |||||  
Db 7141 GACCCACACCCCTTGGCGGGCTGCGTGGAGACAGTACACACACTCCAGTCAAT 7200  
QY |||||  
Db 7201 TCCTGGCTAGGCAATCATCATGTATGCGCCCTTGTGGCAAGGATGATCTGATG 7260  
QY |||||  
Db 7201 TCCTGGCTAGGCAATCATCATGTATGCGCCCTTGTGGCAAGGATGATCTGATG 7260  
QY |||||  
Db 7261 ACTCATTTCTTCTCCATCTCTAGCTCAGGAACAACTTGAAGAGCCCTAGATGTCAG 7320  
QY |||||  
Db 7261 ACTCATTTCTTCTCCATCTCTAGCTCAGGAACAACTTGAAGAGCCCTAGATGTCAG 7320  
QY |||||

QY 7321 ATCTACGGGGCTGTCTTACTTCCATTGAGCCACTTGACTCTCAGATCATTTCAAGACTC 7380  
Db |||||  
QY 7321 ATCTACGGGGCTGTCTTACTTCCATTGAGCCACTTGACTCTCAGATCATTTCAAGACTC 7380  
Db |||||  
QY 7381 CATGCCCTTAGCGCATTTTCTACTCATAGTTACTTCCAGGTGAGATCAATAGGTTGGCT 7440  
Db |||||  
QY 7381 CATGCCCTTAGCGCATTTTCTACTCATAGTTACTTCCAGGTGAGATCAATAGGTTGGCT 7440  
Db |||||  
QY 7441 TCATGCCCTCAGAAACTTTGGGTACCGCCTTGCGAGTCTGAGACATCGGGCCAGAACT 7500  
Db |||||  
QY 7441 TCATGCCCTCAGAAACTTTGGGTACCGCCTTGCGAGTCTGAGACATCGGGCCAGAACT 7500  
Db |||||  
QY 7501 GTTCCGCTTAGGCTACTGTCTCCAGGGGGAGGGCTGCGACTTGTGGCAAGTACTCTTC 7560  
Db |||||  
QY 7501 GTTCCGCTTAGGCTACTGTCTCCAGGGGGAGGGCTGCGACTTGTGGCAAGTACTCTTC 7560  
Db |||||  
QY 7561 AATCGGCGAGTAAGACCAAGCTCAAACTCACTCCATCCCGCTGCGTCCAGTTGGAT 7620  
Db |||||  
QY 7561 AATCGGCGAGTAAGACCAAGCTCAAACTCACTCCATCCCGCTGCGTCCAGTTGGAT 7620  
Db |||||  
QY 7621 TTATCCAGCTGCTGTTGCTGCTTACAGGGGGAGAGATATATCAAGCCCTGTCTCGT 7680  
Db |||||  
QY 7621 TTATCCAGCTGCTGTTGCTGCTTACAGGGGGAGAGATATATCAAGCCCTGTCTCGT 7680  
Db |||||  
QY 7681 GCCCGACCCGCTGTTGCTGCTTACAGGGGGAGAGATATATCAAGCCCTGTCTCGT 7740  
Db |||||  
QY 7681 GCCCGACCCGCTGTTGCTGCTTACAGGGGGAGAGATATATCAAGCCCTGTCTCGT 7740  
Db |||||  
QY 7741 CTACTCCCAACCGATGAAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Db |||||  
QY 7741 CTACTCCCAACCGATGAAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTGTTTT 7800  
Db |||||  
QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db |||||  
QY 7861 TTTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7861 TTTTCTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db |||||  
QY 7921 TAGCTGTGAAGTCCGTTGAGCCGCTGAGAGAGTGTGATAGTGTGATAGTGTGATAGTGTG 7980  
Db |||||  
QY 7921 TAGCTGTGAAGTCCGTTGAGCCGCTGAGAGAGTGTGATAGTGTGATAGTGTGATAGTGTG 7980  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||  
QY 7981 AGATCAAGT 7989  
Db |||||

## RESULT 8

US-10-434-842-6

; Sequence 6, Application US/10434842

; Publication No. US2004005549A1

; GENERAL INFORMATION:

; APPLICANT: Bichko, Vadim

; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY RE

; FILE REFERENCE: 0342/1H395U53

; CURRENT APPLICATION NUMBER: US/10/434,842

; CURRENT FILING DATE: 2003-05-09

; PRIOR APPLICATION NUMBER: US 10/233,307

; PRIOR FILING DATE: 2002-08-28

; PRIOR APPLICATION NUMBER: US 10/005,469

; PRIOR FILING DATE: 2001-11-07

; PRIOR APPLICATION NUMBER: US 60/245,866

; PRIOR FILING DATE: 2000-11-07

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; LENGTH: 7992

; TYPE: DNA

; ORGANISM: Artificial Sequence

; FEATURE:

OTHER INFORMATION: HCV24 subgenomic HCV replicon  
US-10-434-842-6

Query Match	100.0%	Score 7985.8	DB 16	Length 7992	
Best Local Similarity	100.0%	Pred. No. 0			
Matches 7987	Conservative 0	Mismatches 2	Indels 0	Gaps 0	
QY 1	GCACGCCCGGATTTGGGGCGACACTCCACCATAGATCACTCCCTCGTGGAGAACTACTG	60			
DB 1	GCACGCCCGGATTTGGGGCGACACTCCACCATAGATCACTCCCTCGTGGAGAACTACTG	60			
QY 61	TCTTCAGCGAGAAACGCTTAGCCATGCGTGTAGTATGAGTGTCTGTCAGCCTCCAGGAC	120			
DB 61	TCTTCAGCGAGAAACGCTTAGCCATGCGTGTAGTATGAGTGTCTGTCAGCCTCCAGGAC	120			
QY 121	CCCCCTCCCGGAGAGCCATAGTGGTCTCGGAACCGGTGAGTACACCGAAATGGCCAG	180			
DB 121	CCCCCTCCCGGAGAGCCATAGTGGTCTCGGAACCGGTGAGTACACCGAAATGGCCAG	180			
QY 181	GACGACCGGGTCTTTCTTGATCAACCGCTCAATGCTGAGATTTGGCGTGCCTCC	240			
DB 181	GACGACCGGGTCTTTCTTGATCAACCGCTCAATGCTGAGATTTGGCGTGCCTCC	240			
QY 241	CGGAGACTGCTAGCGAGTAGTGTGGTTCGGAAGGCTTGTGTAATGCTGATAGG	300			
DB 241	CGGAGACTGCTAGCGAGTAGTGTGGTTCGGAAGGCTTGTGTAATGCTGATAGG	300			
QY 301	GTGCTTGGAGTGGCCGGGAGGCTCTGACACGTGCACCATGAGCAGCAATCTTAAC	360			
DB 301	GTGCTTGGAGTGGCCGGGAGGCTCTGACACGTGCACCATGAGCAGCAATCTTAAC	360			
QY 361	CTCAAGAAACCAAGGCGCGCCATGATTGAACAGATGGAATGCACGAGGTTCTC	420			
DB 361	CTCAAGAAACCAAGGCGCGCCATGATTGAACAGATGGAATGCACGAGGTTCTC	420			
QY 421	CGGCGGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAGCAATCGGCTGCT	480			
DB 421	CGGCGGCTTGGGTGGAGAGGCTATTCCGCTATGACTGGGCAACAGCAATCGGCTGCT	480			
QY 481	CTGATGCGCGGCTTCCGGGCTGTCAGGCGAGGGGCGCGGTTCTTTTGTCAAGACCG	540			
DB 481	CTGATGCGCGGCTTCCGGGCTGTCAGGCGAGGGGCGCGGTTCTTTTGTCAAGACCG	540			
QY 541	ACCTGTCCGGTCCCTCAATGAACCTGCAGAGCAGGCGCGCTATCGTGGCTGSCCA	600			
DB 541	ACCTGTCCGGTCCCTCAATGAACCTGCAGAGCAGGCGCGCTATCGTGGCTGSCCA	600			
QY 601	CGACGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGC	660			
DB 601	CGACGGCGTTCCTTGGCGAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGACTGGC	660			
QY 661	TGCTATTGGGCGAAGTCCCGGGGAGGATCTCCTGTCTCATCTCACTTGTCTCTGCGGAG	720			
DB 661	TGCTATTGGGCGAAGTCCCGGGGAGGATCTCCTGTCTCATCTCACTTGTCTCTGCGGAG	720			
QY 721	AAGTATCCATCATGGCTGATGCAATGCGGCGGTGCTACGTTGATCCGCGTACCTGCC	780			
DB 721	AAGTATCCATCATGGCTGATGCAATGCGGCGGTGCTACGTTGATCCGCGTACCTGCC	780			
QY 781	CATTTCAGCAACCAAGCGAAACATCGCATGAGCGAGCACTCTCGGATGGAAGCGGTC	840			
DB 781	CATTTCAGCAACCAAGCGAAACATCGCATGAGCGAGCACTCTCGGATGGAAGCGGTC	840			
QY 841	TTGTTCGATCAGGATGATCTGGAGCAAGAGCATCAGGGGCTCGGCCAGCCGAACTGTTCG	900			
DB 841	TTGTTCGATCAGGATGATCTGGAGCAAGAGCATCAGGGGCTCGGCCAGCCGAACTGTTCG	900			
QY 901	CCAGGCTCAAGGCGCGCATCCCGAGCGGAGGATCTCGTGTGACCCATGGCGATGCTCT	960			
DB 901	CCAGGCTCAAGGCGCGCATCCCGAGCGGAGGATCTCGTGTGACCCATGGCGATGCTCT	960			
QY 961	GCTTGGCGAATATCATGGTGGAAATGGCCGCTTTTCTGGAATCATCGACTGTGGCGGC	1020			
DB 961	GCTTGGCGAATATCATGGTGGAAATGGCCGCTTTTCTGGAATCATCGACTGTGGCGGC	1020			

Db	961	GCTTGGCGAATATCATGGTGGAAATGGCCGCTTTTCTGGAATCATCGACTGTGGCGGC	1020
QY	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGAGCTTGGCTACCGGTGATATTTGCTGAAGAGC	1080
Db	1021	TGGGTGTGGCGGACCGCTATCAGGACATAGAGCTTGGCTACCGGTGATATTTGCTGAAGAGC	1080
QY	1081	TTGGCGGGAATGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTGCG	1140
Db	1081	TTGGCGGGAATGGCTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTGCG	1140
QY	1141	AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTCTTAAACAGACCAACAG	1200
Db	1141	AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTCTTAAACAGACCAACAG	1200
QY	1201	GTTTCCCTCTAGCGGATCAATTCGCCCTCTCCCTCCCGCCCTTAAAGTACTGCG	1260
Db	1201	GTTTCCCTCTAGCGGATCAATTCGCCCTCTCCCTCCCGCCCTTAAAGTACTGCG	1260
QY	1261	CGAAGCGCTTGGATTAAGGCGGCTGCGTGTCTTATATGTTATTTTCCACCATATG	1320
Db	1261	CGAAGCGCTTGGATTAAGGCGGCTGCGTGTCTTATATGTTATTTTCCACCATATG	1320
QY	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTCGCCCTCTGTTTGAATGTCGTGAAGAGCA	1380
Db	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAAACCTCGCCCTCTGTTTGAATGTCGTGAAGAGCA	1380
QY	1381	AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGAGCA	1440
Db	1381	AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCGTGAAGAGCA	1440
QY	1441	GTTTCTCTGGAAGCTTCTTGAAGCAAAACAGTCTGTAGCGACCTTTTGACGAGCGG	1500
Db	1441	GTTTCTCTGGAAGCTTCTTGAAGCAAAACAGTCTGTAGCGACCTTTTGACGAGCGG	1500
QY	1501	AACCCCGCTCTGCGGACAGTCTCTCGGCGCAAGCAAGCAGCTGTATAGATACACCT	1560
Db	1501	AACCCCGCTCTGCGGACAGTCTCTCGGCGCAAGCAAGCAGCTGTATAGATACACCT	1560
QY	1561	GCAAGGCGGCAACCCAGTCTGAGTGTGGATGTTGTGGAAAGAGTCAAA	1620
Db	1561	GCAAGGCGGCAACCCAGTCTGAGTGTGGATGTTGTGGAAAGAGTCAAA	1620
QY	1621	TGGCTCTCTCAAGGCTTCAACAGGCGCTGAAGGATGCCAGAGGTACCCCATTTG	1680
Db	1621	TGGCTCTCTCAAGGCTTCAACAGGCGCTGAAGGATGCCAGAGGTACCCCATTTG	1680
QY	1681	ATGGGATCTCATCTGGGCGCTCGGTGCACATGTTTACATGTTTGTAGTCGAGGTTAAA	1740
Db	1681	ATGGGATCTCATCTGGGCGCTCGGTGCACATGTTTACATGTTTGTAGTCGAGGTTAAA	1740
QY	1741	AACGTCTAGGCGGCGGACCGGAGCGTGGTTCCTTTGAAAACACGATAATACC	1800
Db	1741	AACGTCTAGGCGGCGGACCGGAGCGTGGTTCCTTTGAAAACACGATAATACC	1800
QY	1801	ATGGCGCTATTACGGCTACTCTCCACAGACGCGAGGCTCTCTGCTGCATCATCACT	1860
Db	1801	ATGGCGCTATTACGGCTACTCTCCACAGACGCGAGGCTCTCTGCTGCATCATCACT	1860
QY	1861	AGGCTCACAGGCGGAGACAGGACAGGCGGAGGTCCAGTGGTCTCCACCGCA	1920
Db	1861	AGGCTCACAGGCGGAGACAGGACAGGCGGAGGTCCAGTGGTCTCCACCGCA	1920
QY	1921	ACAAATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTGGACTGCTCTATCATGTGTC	1980
Db	1921	ACAAATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTGGACTGCTCTATCATGTGTC	1980
QY	1981	GGCTCAAGACCTTTCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2040
Db	1981	GGCTCAAGACCTTTCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2040
QY	2041	CAGGACCTCGTGGCTGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2100
Db	2041	CAGGACCTCGTGGCTGGCAAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	2100

Qy	2101	GGCAGCTCGACCTTTACTTGGTACAGAGGATGCGATGTCAATTCGCGTGGCCCGCGG	2160	Qy	3181	CAGCGCGAGCGAGGACTGGTAGGCGCAGGATGGCAATTTACAGGTTGTGTGACTCAGGA	3240
Db	2101	GGCAGCTCGACCTTTACTTGGTACAGAGGATGCGATGTCAATTCGCGTGGCCCGCGG	2160	Db	3181	CAGCGCGAGCGAGGACTGGTAGGCGCAGGATGGCAATTTACAGGTTGTGTGACTCAGGA	3240
Qy	2161	GGCGACAGAGGGGAGGACTACTCTCCCGCAGGCGCGCTCTCTACTTGAAGGCTCTTCG	2220	Qy	3241	GAAACGCCCTCGGGCATGTTGATTCCTGGTTCTGTGGAGTGTATGACGCGGGCTGT	3300
Db	2161	GGCGACAGAGGGGAGGACTACTCTCCCGCAGGCGCGCTCTCTACTTGAAGGCTCTTCG	2220	Db	3241	GAAACGCCCTCGGGCATGTTGATTCCTGGTTCTGTGGAGTGTATGACGCGGGCTGT	3300
Qy	2221	GGCGGTCCACTCTCTGCGCCCTCGGGGACAGCTGTGGGCATCTTTTCGGGCTCCGCTGTGC	2280	Qy	3301	GCTTGGTAGAGCTCACGCCCGCGAGACCTCAGTTAGTTTTCGGGCTTACCTAAACACA	3360
Db	2221	GGCGGTCCACTCTCTGCGCCCTCGGGGACAGCTGTGGGCATCTTTTCGGGCTCCGCTGTGC	2280	Db	3301	GCTTGGTAGAGCTCACGCCCGCGAGACCTCAGTTAGTTTTCGGGCTTACCTAAACACA	3360
Qy	2281	ACCCGAGGGGTGCGAAGCGGTGTGACTTTGTACCGTCTGAGTCTATGGAACCACTATG	2340	Qy	3361	CCAGGGTTGCCCGTCTGCAGGACCATCTGGAGTTCTGGGAGAGCTCTTTACAGGCTTC	3420
Db	2281	ACCCGAGGGGTGCGAAGCGGTGTGACTTTGTACCGTCTGAGTCTATGGAACCACTATG	2340	Db	3361	CCAGGGTTGCCCGTCTGCAGGACCATCTGGAGTTCTGGGAGAGCTCTTTACAGGCTTC	3420
Qy	2341	CGGTCCCAGGTCTTCAAGGACAACTCGTCCCTTCGGGCGGTACCGGAGACATTCAGGTG	2400	Qy	3421	ACCCACATAGACGCCCATTTCTTGTCCAGACTTAAGCAGGAGGAGAGCAACTTCCCTAC	3480
Db	2341	CGGTCCCAGGTCTTCAAGGACAACTCGTCCCTTCGGGCGGTACCGGAGACATTCAGGTG	2400	Db	3421	ACCCACATAGACGCCCATTTCTTGTCCAGACTTAAGCAGGAGGAGAGCAACTTCCCTAC	3480
Qy	2401	GCCCATCTACGCCCTTACTGTAAGCGCAAGAGCACTAAGGTGCGGCTGCTGTATGCA	2460	Qy	3481	CTGGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCGTGGAC	3540
Db	2401	GCCCATCTACGCCCTTACTGTAAGCGCAAGAGCACTAAGGTGCGGCTGCTGTATGCA	2460	Db	3481	CTGGTAGCATACAGGCTACGGTGTGCGCCAGGCTCAGGCTCCACCTCCATCGTGGAC	3540
Qy	2461	GCCAAAGGTATAGGTGCTTGTCTGAACCCGTTCGGTCCGCGCCACCTTAGGTTTCGGG	2520	Qy	3541	CAAAATGTGAAGTGTCTCATACGGCTAAAGCTACGCTGCAAGGCGCAACGCGCTGCTG	3600
Db	2461	GCCAAAGGTATAGGTGCTTGTCTGAACCCGTTCGGTCCGCGCCACCTTAGGTTTCGGG	2520	Db	3541	CAAAATGTGAAGTGTCTCATACGGCTAAAGCTACGCTGCAAGGCGCAACGCGCTGCTG	3600
Qy	2521	GGGTATCTTAAGGACATGTAATCGACCTTAACATCAGAACCGGGGTAAAGACCATC	2580	Qy	3601	TATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACCCATACCAANTACATC	3660
Db	2521	GGGTATCTTAAGGACATGTAATCGACCTTAACATCAGAACCGGGGTAAAGACCATC	2580	Db	3601	TATAGGCTGGGAGCGGTTCAAAACGAGGTTACTACACACACCCATACCAANTACATC	3660
Qy	2581	ACCACGGGTGCCCCATCAGTACTCCACTATGGAAGTTCTTCCGACGGTGTGTTC	2640	Qy	3661	ATGGCATGATGTGCGGCTCACCTGGAGTGTCTCACGAGCACCTGGGTGTGTAGGCGGA	3720
Db	2581	ACCACGGGTGCCCCATCAGTACTCCACTATGGAAGTTCTTCCGACGGTGTGTTC	2640	Db	3661	ATGGCATGATGTGCGGCTCACCTGGAGTGTCTCACGAGCACCTGGGTGTGTAGGCGGA	3720
Qy	2641	TTTGGGGGCGCTATGACATCAATATGTGATGAGTGCCACTCAACTGACTCGACACT	2700	Qy	3721	GTCTAGCAGCTCTGCGCGCTATTCCTGCAACAGGAGCGTGTGTATTTGTGGGAGG	3780
Db	2641	TTTGGGGGCGCTATGACATCAATATGTGATGAGTGCCACTCAACTGACTCGACACT	2700	Db	3721	GTCTAGCAGCTCTGCGCGCTATTCCTGCAACAGGAGCGTGTGTATTTGTGGGAGG	3780
Qy	2701	ATCTGGGATCGGACAGTCTTGGACCAAGCGGAGAGCGCTGAGCGGCGACTCTGCTG	2760	Qy	3781	ATCATCTTGTCCGGAAGCGGCGCATCTCCGACAGGGAAGTCTTTTACCGGGAGTTC	3840
Db	2701	ATCTGGGATCGGACAGTCTTGGACCAAGCGGAGAGCGCTGAGCGGCGACTCTGCTG	2760	Db	3781	ATCATCTTGTCCGGAAGCGGCGCATCTCCGACAGGGAAGTCTTTTACCGGGAGTTC	3840
Qy	2761	CTGCCACCGCTACGGCTCCGGATCGGTCAACCGTGCACATCCAAACATCGAGGAGTG	2820	Qy	3841	GATGAGATGGAAGAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGACAGCTGCC	3900
Db	2761	CTGCCACCGCTACGGCTCCGGATCGGTCAACCGTGCACATCCAAACATCGAGGAGTG	2820	Db	3841	GATGAGATGGAAGAGTGCCTCACACCTCCCTTACATCGAACAGGGAATGACAGCTGCC	3900
Qy	2821	GCTCTGTCCAGACTCGGAAATCCCTTTATGCGAAGCCATCCCGGAGGAGGAGTTC	2880	Qy	3901	GAAACAAATTCAAAACAGAGGCAATCGGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGCT	3960
Db	2821	GCTCTGTCCAGACTCGGAAATCCCTTTATGCGAAGCCATCCCGGAGGAGGAGTTC	2880	Db	3901	GAAACAAATTCAAAACAGAGGCAATCGGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGCT	3960
Qy	2881	AAGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2940	Qy	3961	GCTGCTCCCGTGGTGAATCCAAAGTGGCGGACCTCGAAGGCTTCTGGGCGAAGCATATG	4020
Db	2881	AAGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	2940	Db	3961	GCTGCTCCCGTGGTGAATCCAAAGTGGCGGACCTCGAAGGCTTCTGGGCGAAGCATATG	4020
Qy	2941	AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGCGCTTGTATTCGCTC	3000	Qy	4021	TGCAATTTTCAATCAGCGGATACAAATTTTAGCAGGCTTGTCCACTCTGCTTGGCAACCCC	4080
Db	2941	AAGCTGTCCGGCTCGGACTCAATGTGTAGCATATTAACGGGCGCTTGTATTCGCTC	3000	Db	4021	TGCAATTTTCAATCAGCGGATACAAATTTTAGCAGGCTTGTCCACTCTGCTTGGCAACCCC	4080
Qy	3001	ATACCAACTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3060	Qy	4081	GCGATAGCACTGATGGCATTCACAGCTCTTATACAGGCGGCTTCAACCCACAT	4140
Db	3001	ATACCAACTAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG	3060	Db	4081	GCGATAGCACTGATGGCATTCACAGCTCTTATACAGGCGGCTTCAACCCACAT	4140
Qy	3061	GGGATTTGACTCAGTGATCGACTGCAATACATGTGTCAACCCAGAGAGGAGGAGGAGGAGG	3120	Qy	4141	ACCTCTCTGTTTAACTCTTGGGGGATGGTGGCGCCCACTTGTCTCTCCAGGCT	4200
Db	3061	GGGATTTGACTCAGTGATCGACTGCAATACATGTGTCAACCCAGAGAGGAGGAGGAGGAGG	3120	Db	4141	ACCTCTCTGTTTAACTCTTGGGGGATGGTGGCGCCCACTTGTCTCTCCAGGCT	4200
Qy	3121	CTGACCGGAGCTTCACTATGAGACGAGCGGCTGCGCAACGAGCGGCTGTACGCTCG	3180	Qy	4201	GCTTCTCTTTCGTAGGCGCGGATCGCTGGAGGCGCTGTGTGGCAGCATAGGCTTGGG	4260
Db	3121	CTGACCGGAGCTTCACTATGAGACGAGCGGCTGCGCAACGAGCGGCTGTACGCTCG	3180	Db	4201	GCTTCTCTTTCGTAGGCGCGGATCGCTGGAGGCGCTGTGTGGCAGCATAGGCTTGGG	4260
				Qy	4261	AAGGTGCTTGTGGATATTTTGGCAGGTTTGGACAGGCGGCGGCGGCTCGTGCC	4320



Db 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGAGCAGGGGTGGCAGGCGCTCGTGGCC 4320  
Qy |||||  
Db 4321 TTTAAGGTGCTATGAGCGGCGAGATGCCCTCCACGGAGACCTGGTTAA CTTACTCCCTGCT 4380  
Qy |||||  
Db 4321 TTTAAGGTGCTATGAGCGGCGAGATGCCCTCCACGGAGACCTGGTTAA CTTACTCCCTGCT 4380  
Qy |||||  
Db 4381 APTCTCTCCCTCGGGGCCCTAGTCGTGCGGGCTGTTGGCGCAGCGATACGTGGTGGGCAC 4440  
Qy |||||  
Db 4381 ATCTCTCTCCCTCGGGGCCCTAGTCGTGCGGGCTGTTGGCGCAGCGATACGTGGTGGGCAC 4440  
Qy |||||  
Db 4441 GTGGGCCACAGGGGAGGGGGCTGTGCACTGATGATGAACCGGCTGATAGCGTTCCGCTCGGG 4500  
Qy |||||  
Db 4441 GTGGGCCACAGGGGAGGGGGCTGTGCACTGATGATGAACCGGCTGATAGCGTTCCGCTCGGG 4500  
Qy |||||  
Db 4501 GGTAAACACAGCTCTCCCGCAGCAGCTATGCTGCTGAGAGGAGCGTGCAGAGCTGTCACT 4560  
Qy |||||  
Db 4501 GGTAAACACAGCTCTCCCGCAGCAGCTATGCTGCTGAGAGGAGCGTGCAGAGCTGTCACT 4560  
Qy |||||  
Db 4561 CAGATCTCTCTAGTCTTTACCATCACTCAGCTGCTGAAGAGCTTACCAGTGGATCAAC 4620  
Qy |||||  
Db 4561 CAGATCTCTCTAGTCTTTACCATCACTCAGCTGCTGAAGAGCTTACCAGTGGATCAAC 4620  
Qy |||||  
Db 4621 GAGGACTGCTCCAGCGCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
Qy |||||  
Db 4621 GAGGACTGCTCCAGCGCATGCTCCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC 4680  
Qy |||||  
Db 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCTCCGCGGATTTGGCGGGA 4740  
Qy |||||  
Db 4681 ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAGTCTCCGCGGATTTGGCGGGA 4740  
Qy |||||  
Db 4741 GTCCCTCTCTCATGTCACAGTGGGTACAAAGGAGTCTGGCGGGGCGAGCGGATCATG 4800  
Qy |||||  
Db 4741 GTCCCTCTCTCATGTCACAGTGGGTACAAAGGAGTCTGGCGGGGCGAGCGGATCATG 4800  
Qy |||||  
Db 4801 CAACACACCTGCTCCATGTGAGGACAGATCAACCGGACATGTTGAAAAACGGTTCCATGAGG 4860  
Qy |||||  
Db 4801 CAACACACCTGCTCCATGTGAGGACAGATCAACCGGACATGTTGAAAAACGGTTCCATGAGG 4860  
Qy |||||  
Db 4861 ATCTGTGGGCTTAGGACTCTAGTAAACAGCTGGGATGAAACATTTCCCATTAACCGGTAC 4920  
Qy |||||  
Db 4861 ATCTGTGGGCTTAGGACTCTAGTAAACAGCTGGGATGAAACATTTCCCATTAACCGGTAC 4920  
Qy |||||  
Db 4921 ACCACGGGCCCCCTGACGCCCTCCCGCGCGCAAAATTTATCTAGGCGCTGTGGCGGGTG 4980  
Qy |||||  
Db 4921 ACCACGGGCCCCCTGACGCCCTCCCGCGCGCAAAATTTATCTAGGCGCTGTGGCGGGTG 4980  
Qy |||||  
Db 4981 GCTGCTGAGGAGTACGTGGAGGTTACCGGGTGGGGGATTTCCACTACGTGACGGGCATG 5040  
Qy |||||  
Db 4981 GCTGCTGAGGAGTACGTGGAGGTTACCGGGTGGGGGATTTCCACTACGTGACGGGCATG 5040  
Qy |||||  
Db 5041 ACCACTGACAACTGAAGTGCCTGTGTCAGGTTCCGGCCCCCGAAATTTCTTCAAGAGTG 5100  
Qy |||||  
Db 5041 ACCACTGACAACTGAAGTGCCTGTGTCAGGTTCCGGCCCCCGAAATTTCTTCAAGAGTG 5100  
Qy |||||  
Db 5101 GATGGGGTGGGTTGACAGGTACGCTCCAGCGTGAACCCCTCTACCGGAGGAGGTC 5160  
Qy |||||  
Db 5101 GATGGGGTGGGTTGACAGGTACGCTCCAGCGTGAACCCCTCTACCGGAGGAGGTC 5160  
Qy |||||  
Db 5161 ACATTCCTGTGGGTCAATCAATCACTGTTGGGTACAGCTCCCATCGAGCCCGGAA 5220  
Qy |||||  
Db 5161 ACATTCCTGTGGGTCAATCAATCACTGTTGGGTACAGCTCCCATCGAGCCCGGAA 5220  
Qy |||||  
Db 5221 CCGGAGTACAGTGTCACTTCCATGCTCAGGACCCCTCCACATTAGCGGAGAGC 5280  
Qy |||||  
Db 5221 CCGGAGTACAGTGTCACTTCCATGCTCAGGACCCCTCCACATTAGCGGAGAGC 5280  
Qy |||||  
Db 5281 GCTAAGCGTAGGCTGGCAGGGGATCTCCCGCTCTTGGCGAGCTATCAGGTATCCAG 5340  
Qy |||||  
Db 5281 GCTAAGCGTAGGCTGGCAGGGGATCTCCCGCTCTTGGCGAGCTATCAGGTATCCAG 5340  
Qy |||||  
Db 5341 CTGCTGCGCTTCTTGAAGGCAACATGACCTACCGCTCATGACTCCCGGAGCTGAC 5400  
Qy |||||

Db 5341 CTGCTGCGCTTCTTGAAGGCAACATGCACTACCCGTCATGACTCCCGGACGCTGAC 5400  
Qy |||||  
Db 5401 CTTCATCGAGGCCCAACCTCTCTGTGGCGGAGAGATGGGGCGGAAACATACCCCGCTGGAG 5460  
Qy |||||  
Db 5401 CTTCATCGAGGCCCAACCTCTCTGTGGCGGAGAGATGGGGCGGAAACATACCCCGCTGGAG 5460  
Qy |||||  
Db 5461 TCAGAAAATAGGTAGTAAATTTTGGACTCTTTCGAGCGCTCCAAGCGGAGGATGAG 5520  
Qy |||||  
Db 5461 TCAGAAAATAGGTAGTAAATTTTGGACTCTTTCGAGCGCTCCAAGCGGAGGATGAG 5520  
Qy |||||  
Db 5521 AGGGAAGTATCCGTTCCCGCGGAGATCTGCGGAGGTCAGGAAATTCCTTCGAGGATG 5580  
Qy |||||  
Db 5521 AGGGAAGTATCCGTTCCCGCGGAGATCTGCGGAGGTCAGGAAATTCCTTCGAGGATG 5580  
Qy |||||  
Db 5581 CCCATATGGGACGACCCCGGATTAACAACCTCCACTGTTTAGAGTCTCTGGAAGGACCCGAC 5640  
Qy |||||  
Db 5581 CCCATATGGGACGACCCCGGATTAACAACCTCCACTGTTTAGAGTCTCTGGAAGGACCCGAC 5640  
Qy |||||  
Db 5641 TAGCTCCCTCAGTGGTACAGGGGTGTCATGTCGCTGCCAAGGCCCTCCGATACCA 5700  
Qy |||||  
Db 5641 TAGCTCCCTCAGTGGTACAGGGGTGTCATGTCGCTGCCAAGGCCCTCCGATACCA 5700  
Qy |||||  
Db 5701 CCTCCACGAGGAAAGAGACGGTTGCTCTGTGTCAGAACTACCGTGTCTTCTGCTTGGG 5760  
Qy |||||  
Db 5701 CCTCCACGAGGAAAGAGACGGTTGCTCTGTGTCAGAACTACCGTGTCTTCTGCTTGGG 5760  
Qy |||||  
Db 5761 GAGCTGCCCCAACAAGACCTTCGGCAGCTCCGAATCGTCGGCGCTCGACAGCGGACCGCA 5820  
Qy |||||  
Db 5761 GAGCTGCCCCAACAAGACCTTCGGCAGCTCCGAATCGTCGGCGCTCGACAGCGGACCGCA 5820  
Qy |||||  
Db 5821 ACAGCCCTCTCTGACAGGCTTCGACAGCCCTCGACAGCGGCGCTCGACAGCGGACCGCA 5880  
Qy |||||  
Db 5821 ACAGCCCTCTCTGACAGGCTTCGACAGCCCTCGACAGCGGCGCTCGACAGCGGACCGCA 5880  
Qy |||||  
Db 5881 TCCTCCATGCCCCCTTTGAGGGGAGCGCGGGATCCCGATCTCTCAGCGAGCGGTCTTGG 5940  
Qy |||||  
Db 5881 TCCTCCATGCCCCCTTTGAGGGGAGCGCGGGATCCCGATCTCTCAGCGAGCGGTCTTGG 5940  
Qy |||||  
Db 5941 TCTACCGTAAGCAGGAGGCTAGTGAGGACGTCTCTGTCTGTCTGATGCTCTACACATGG 6000  
Qy |||||  
Db 5941 TCTACCGTAAGCAGGAGGCTAGTGAGGACGTCTCTGTCTGTCTGATGCTCTACACATGG 6000  
Qy |||||  
Db 6001 ACAGGCGCCCTGATCAGCCATCGCTGCGGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
Qy |||||  
Db 6001 ACAGGCGCCCTGATCAGCCATCGCTGCGGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
Qy |||||  
Db 6061 AGCAACTCTTTGCTCCGTACCACTTGGTCTATGCTATGCTACCACTCTCGAGGCGAAGC 6120  
Qy |||||  
Db 6061 AGCAACTCTTTGCTCCGTACCACTTGGTCTATGCTATGCTACCACTCTCGAGCGCAAGC 6120  
Qy |||||  
Db 6121 CTGCGGACAGGAGGTCACCTTTGACAGACTGCGAGTCTCTGACGACCACTACCGGAC 6180  
Qy |||||  
Db 6121 CTGCGGACAGGAGGTCACCTTTGACAGACTGCGAGTCTCTGACGACCACTACCGGAC 6180  
Qy |||||  
Db 6181 GTGCTCAAGAGATGAAGGCGAAGCGTCCACAGTTAAGGCTAAACTCTATCCGTTGGAG 6240  
Qy |||||  
Db 6181 GTGCTCAAGAGATGAAGGCGAAGCGTCCACAGTTAAGGCTAAACTCTATCCGTTGGAG 6240  
Qy |||||  
Db 6241 GAAGCTGTAAAGTGAAGCGGCGGACCACTTGGCCAGATCTAAATTTGGCTATGAGGCAAG 6300  
Qy |||||  
Db 6241 GAAGCTGTAAAGTGAAGCGGCGGACCACTTGGCCAGATCTAAATTTGGCTATGAGGCAAG 6300  
Qy |||||  
Db 6301 GAGCTCCGGAACCTATTCAGCAAGGCGGTTAACCACTCCGCTCCGCTGAGGAGGACTTG 6360  
Qy |||||  
Db 6301 GAGCTCCGGAACCTATTCAGCAAGGCGGTTAACCACTCCGCTCCGCTGAGGAGGACTTG 6360  
Qy |||||  
Db 6361 CTGGAAGACACTGAGACCAATTTGACACCACTCATGCGAAGAAATGAGGTTTCTGC 6420  
Qy |||||  
Db 6361 CTGGAAGACACTGAGACCAATTTGACACCACTCATGCGAAGAAATGAGGTTTCTGC 6420  
Qy |||||  
Db 6421 GTCCACACAGAGAGGGGGCGGAGCGGCTTATCGTATTCAGATTTGGGG 6480  
Qy |||||  
Db 6421 GTCCACACAGAGAGGGGGCGGAGCGGCTTATCGTATTCAGATTTGGGG 6480



Db 181 GACGACCGGTCCTTTCTTGGATCAACCGCTCAATGCGCTGAGATTGGGCGTCCGCC 240  
QY 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGCCCTTGTGTGACTGCTGATAGG 300  
Db 241 GCGAGACTGTAGCCGAGTAGTGTGGGTGCGGAAAGCCCTTGTGTGACTGCTGATAGG 300  
QY 301 GTGCTTGGAGTGGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360  
Db 301 GTGCTTGGAGTGGCCCGGAGGTCTCGTAGACCGTGCACCATGAGCAGCAATCCTAAAC 360  
QY 361 CTCAAAGAAAACAAAGGGCGCCCATGATGAAACAAGATGGAATGCAACGAGTCTTC 420  
Db 361 CTCAAAGAAAACAAAGGGCGCCCATGATGAAACAAGATGGAATGCAACGAGTCTTC 420  
QY 421 CGGCCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCAACAGACAATCGGCTGCT 480  
Db 421 CGGCCGCTTGGGTGAGAGGCTATTCGGCTATGACTGGGCAACAGACAATCGGCTGCT 480  
QY 481 CTGATGCGCGCGTGTTCGGGCTGTGAGCGAGGGGCGCCGCTTCTTTTGTCAAGACCG 540  
Db 481 CTGATGCGCGCGTGTTCGGGCTGTGAGCGAGGGGCGCCGCTTCTTTTGTCAAGACCG 540  
QY 541 ACCTGTCCGGTCCCTGAATGAACTGCAGGACGAGGCAAGCGGCTATCGTGGCTGGCCA 600  
Db 541 ACCTGTCCGGTCCCTGAATGAACTGCAGGACGAGGCAAGCGGCTATCGTGGCTGGCCA 600  
QY 601 CGAGCGGCTTCCCTGCGACGTGTCTGCAAGTGTCTGCAAGCGGGAAGGACTGGC 660  
Db 601 CGAGCGGCTTCCCTGCGACGTGTCTGCAAGTGTCTGCAAGCGGGAAGGACTGGC 660  
QY 661 TGTCTATTTGGGCGAAGTGCAGGCGGAGGATCTCTGTGATCTCACTTGTCTCCCGGAGA 720  
Db 661 TGTCTATTTGGGCGAAGTGCAGGCGGAGGATCTCTGTGATCTCACTTGTCTCCCGGAGA 720  
QY 721 AAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATAGCTTTGATCCGGCTACCTGCC 780  
Db 721 AAGTATCCATCATGGCTGATGCAATGCGGCGGCTGCATAGCTTTGATCCGGCTACCTGCC 780  
QY 781 CATTCGACCAACAGCGAATCATCGATCGAGGACGACGATCTCGGATGGAAGCGGTC 840  
Db 781 CATTCGACCAACAGCGAATCATCGATCGAGGACGACGATCTCGGATGGAAGCGGTC 840  
QY 841 TTGTGATCAGGATGATCTGGAAGAGGATCAGGGGCTCGGCCAGCGAATGTTTCG 900  
Db 841 TTGTGATCAGGATGATCTGGAAGAGGATCAGGGGCTCGGCCAGCGAATGTTTCG 900  
QY 901 CAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGTGTGACCCATGGCGATGCT 960  
Db 901 CAGGCTCAAGGCGCGCATGCCGACGCGGAGGATCTCGTGTGACCCATGGCGATGCT 960  
QY 961 GCTTGGCGAATCATGCTGGAATAAGCGCTTTCTGGAATCATCGATGGAAGCGG 1020  
Db 961 GCTTGGCGAATCATGCTGGAATAAGCGCTTTCTGGAATCATCGATGGAAGCGG 1020  
QY 1021 TGGGTGTGGCGGACCGCTATCAGACATAGCTGTGGCTACCGGTATCGGCTCCCGATTCG 1080  
Db 1021 TGGGTGTGGCGGACCGCTATCAGACATAGCTGTGGCTACCGGTATCGGCTCCCGATTCG 1080  
QY 1081 TTGGCGGCAATGGGTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCG 1140  
Db 1081 TTGGCGGCAATGGGTGACCGCTTCTCGTGTCTTACGGTATCGCGCTCCCGATTCG 1140  
QY 1141 AGCGCATCGCTTCTATCGCTTCTTGAACAGTCTTCTGAGTTTAAACAGACCAACG 1200  
Db 1141 AGCGCATCGCTTCTATCGCTTCTTGAACAGTCTTCTGAGTTTAAACAGACCAACG 1200  
QY 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTTCCCTCCCGCTTAAACAGTCTTGGC 1260  
Db 1201 GTTTCCTCTAGCGGATCAATTCGCGCTCTTCCCTCCCGCTTAAACAGTCTTGGC 1260  
QY 1261 CGAAGCGCTTGAATTAAGCGCGGTGTGGTGTCTATATGTTATTTCCACCATATTG 1320

Db 1261 CGAAGCGCTTGAATTAAGCGCGGTGTGGTGTCTATATGTTATTTCCACCATATTG 1320  
QY 1321 CGGTCTTTTGGCAATGTGAGGCCCCGAAAACTGCGCCCTGTCTTCTTTCAGGACATTCCT 1380  
Db 1321 CGGTCTTTTGGCAATGTGAGGCCCCGAAAACTGCGCCCTGTCTTCTTTCAGGACATTCCT 1380  
QY 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCTGTAAGGACA 1440  
Db 1381 AGGGGTCTTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCTGTAAGGACA 1440  
QY 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACAGCTCTGAGCAGCCCTTTCAGGCGAGGG 1500  
Db 1441 GTTCTCTGGAAGCTTCTTGAAGACAAACAGCTCTGAGCAGCCCTTTCAGGCGAGGG 1500  
QY 1501 AACCCCCACCTGCGCACAGGTGCTCTGCGGCCAAAAAGCCACGTTATAAGATACACT 1560  
Db 1501 AACCCCCACCTGCGCACAGGTGCTCTGCGGCCAAAAAGCCACGTTATAAGATACACT 1560  
QY 1561 GCAAGGCGGCACAAACCCAGTGCACCTTGTGAGTTGGAATGTTGGAAGAGTCAAA 1620  
Db 1561 GCAAGGCGGCACAAACCCAGTGCACCTTGTGAGTTGGAATGTTGGAAGAGTCAAA 1620  
QY 1621 TGGCTCTCTCAAGCTATTCAACAAGGCGCTGAAGGATGCCAGAGGTACCCCATTTG 1680  
Db 1621 TGGCTCTCTCAAGCTATTCAACAAGGCGCTGAAGGATGCCAGAGGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGTTAAA 1740  
Db 1681 ATGGGATCTGATCTGGGCGCTCGGTGCACATGCTTTACATGTTTGTAGTCGAGTTAAA 1740  
QY 1741 AACGTCTAGGCCCCCGGAAACACGCGGACGTTGTTTCCCTTTGAAAAACAGATAACC 1800  
Db 1741 AACGTCTAGGCCCCCGGAAACACGCGGACGTTGTTTCCCTTTGAAAAACAGATAACC 1800  
QY 1801 ATGGCGCTATTACGGCTACTCCACAGACGCGAGGCGCTACTTGGCTGCATCATCT 1860  
Db 1801 ATGGCGCTATTACGGCTACTCCACAGACGCGAGGCGCTACTTGGCTGCATCATCT 1860  
QY 1861 AGCCTCAAGCGCGGACAGGAAACAGGTGAGGGGAGGTCCAAGTGGTCTCCACCGCA 1920  
Db 1861 AGCCTCAAGCGCGGACAGGAAACAGGTGAGGGGAGGTCCAAGTGGTCTCCACCGCA 1920  
QY 1921 ACACATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTTGGACTGTCTATCATGTGCC 1980  
Db 1921 ACACATCTTTCTGCGGACCTGCGTCAATGGCGTGTGTTGGACTGTCTATCATGTGCC 1980  
QY 1981 GGTCTAAAGACCTTGC CGGCCCAAGGGGCCAATCACCCAAATGTACACCAATGTGGAC 2040  
Db 1981 GGTCTAAAGACCTTGC CGGCCCAAGGGGCCAATCACCCAAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACCTCGTGGCTGGCAAGCGCCCCCGGCGCGTTCCTTTGACACCATGCACCTGC 2100  
Db 2041 CAGGACCTCGTGGCTGGCAAGCGCCCCCGGCGCGTTCCTTTGACACCATGCACCTGC 2100  
QY 2101 GGCAGCTCGAATCTTTACTTGTGTGACAGCATGCGATGTCATTCGCGTGC CGCGCGG 2160  
Db 2101 GGCAGCTCGAATCTTTACTTGTGTGACAGCATGCGATGTCATTCGCGTGC CGCGCGG 2160  
QY 2161 GCGGACAGCAGGAGGAGCTACTCTCCCGAGCGCTCTCTCTACTTTGAAGGCTCTTCG 2220  
Db 2161 GCGGACAGCAGGAGGAGCTACTCTCCCGAGCGCTCTCTCTACTTTGAAGGCTCTTCG 2220  
QY 2221 GCGGTCTCACTGTCTGCGCCCTCGGGGACAGCTGTGGGCTCTTTTCGGGCTCGCGGTGC 2280  
Db 2221 GCGGTCTCACTGTCTGCGCCCTCGGGGACAGCTGTGGGCTCTTTTCGGGCTCGCGGTGC 2280  
QY 2281 ACCGAGGGGTTCGGAAGCGGTGGAATTTGTACCGTGCAGTCTATGGAACCACTATG 2340  
Db 2281 ACCGAGGGGTTCGGAAGCGGTGGAATTTGTACCGTGCAGTCTATGGAACCACTATG 2340  
QY 2341 CGGTCCCGGCTTTCAGGACAACTCTGTCCTTCGGCGGTAACCGGAGCATTTCCAGGTG 2400  
Db 2341 CGGTCCCGGCTTTCAGGACAACTCTGTCCTTCGGCGGTAACCGGAGCATTTCCAGGTG 2400



Db	4561		CAGATCCTCTCTAGTCTTTACCACTACTCAGCTGCTGAAGAGGCTTCACCACTGGATCAAC	4620
QY	4621		GAGGACTGCTCCACGCCATGTCTCCGGCTCTGGCTTAAGAGATGTTTGGATTTGGATATGC	4680
Db	4621		GAGGACTGCTCCACGCCATGCTCCGGCTCTGGCTTAAGAGATGTTTGGATTTGGATATGC	4680
QY	4681		ACGGTGTGACTGATTTCAAGACTTGGGCTCCAGTCCAAAGCTCTCTCGCGCATTCGCGGA	4740
Db	4681		ACGGTGTGACTGATTTCAAGACTTGGGCTCCAGTCCAAAGCTCTCTCGCGCATTCGCGGA	4740
QY	4741		GTCCCTCTCTCTCATGTCAACGTGGGTACAAAGGAGTCTGGCGGGGCGACGGATCATG	4800
Db	4741		GTCCCTCTCTCTCATGTCAACGTGGGTACAAAGGAGTCTGGCGGGGCGACGGATCATG	4800
QY	4801		CAAAACCACTCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCATCAGG	4860
Db	4801		CAAAACCACTCCCATGTGGAGCAGATCACCGGACATGTGAAAAACGGTTCATCAGG	4860
QY	4861		ATCGTGGGCTTAGGACTGTAGTAAACAGTGGGATGAAACATTTCCCATTAACCGGTAC	4920
Db	4861		ATCGTGGGCTTAGGACTGTAGTAAACAGTGGGATGAAACATTTCCCATTAACCGGTAC	4920
QY	4921		ACCAAGGCGCCCTGACGCCCTCCCGGCGCCAAATTTATTTAGGGCGCTGTGGCGGGTG	4980
Db	4921		ACCAAGGCGCCCTGACGCCCTCCCGGCGCCAAATTTATTTAGGGCGCTGTGGCGGGTG	4980
QY	4981		GCTGCTGAGGAGTACGTGGAGTTTACCGGGTGGGGATTTTCCACTACGTGACGGGCATG	5040
Db	4981		GCTGCTGAGGAGTACGTGGAGTTTACCGGGTGGGGATTTTCCACTACGTGACGGGCATG	5040
QY	5041		ACCACTGACAACTGAAAGTGCCTGTCAAGTTCCTGGCCCCGAAATTTCTTCACAGAAAGTG	5100
Db	5041		ACCACTGACAACTGAAAGTGCCTGTCAAGTTCCTGGCCCCGAAATTTCTTCACAGAAAGTG	5100
QY	5101		GATGGGTGCGTTGACAGGTACGCTCCAGCGTGCAAAACCTCTCTACGGGAGGAGTTC	5160
Db	5101		GATGGGTGCGTTGACAGGTACGCTCCAGCGTGCAAAACCTCTCTACGGGAGGAGTTC	5160
QY	5161		ACATTCCTGTGCGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCCCGAA	5220
Db	5161		ACATTCCTGTGCGGCTCAATCAATACCTGTTGGGTACAGCTCCCATGCGAGCCCGAA	5220
QY	5221		CCGAGCTAGCAGTGCTCATTCATGCTCACCGACCCCTCCACATTAACGGCGGAGACG	5280
Db	5221		CCGAGCTAGCAGTGCTCATTCATGCTCACCGACCCCTCCACATTAACGGCGGAGACG	5280
QY	5281		GCTAAGCGTAGGTGGCGAGGGATCTCCCCCTCTGGCCAGTCTACGTAGCCAG	5340
Db	5281		GCTAAGCTTAGGTGGCGAGGGATCTCCCCCTCTGGCCAGTCTACGTAGCCAG	5340
QY	5341		CTGTCTGCGCTTCTTGGAAGCAACATGACATACCCGTCTAGCTCCCCCGACGCTGAC	5400
Db	5341		CTGTCTGCGCTTCTTGGAAGCAACATGACATACCCGTCTAGCTCCCCCGACGCTGAC	5400
QY	5401		CTCATCGAGCCAACTCTGTGGCGGAGAGATGGGCGGAAACATACCCCGTGGAG	5460
Db	5401		CTCATCGAGCCAACTCTGTGGCGGAGAGATGGGCGGAAACATACCCCGTGGAG	5460
QY	5461		TCAGAAAAAAGGTAGTAAATTTTGGACTCTTTCGAGCCGCTCCAGCGGAGAGATGAG	5520
Db	5461		TCAGAAAAAAGGTAGTAAATTTTGGACTCTTTCGAGCCGCTCCAGCGGAGAGATGAG	5520
QY	5521		AGGGAAGTATCCGTTCCGGCGGAGATCTCTCGGAGGTCCAGGAAATTCCTCGACGGATG	5580
Db	5521		AGGGAAGTATCCGTTCCGGCGGAGATCTCTCGGAGGTCCAGGAAATTCCTCGACGGATG	5580
QY	5581		CCCATATGGGACGCGCCGGATTAAACCTCCACTGTTAGAGTCTCTGGAAGGACCCGGAC	5640
Db	5581		CCCATATGGGACGCGCCGGATTAAACCTCCACTGTTAGAGTCTCTGGAAGGACCCGGAC	5640
QY	5641		TACGTCCTCCAGTGGTACACGGTGTTCATTGCGCGCTGCCAAGGCCCTCCGATACCA	5700

Db	5641	TACGTCCTCCAGTGGTACACGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA	5700
Qy	5701	CCTCCACGGAGGAGAGACGGTTGTCTCTGTAGAACTACCGTGTCTTCTGCTTGGCG	5760
Db	5701	CCTCCACGGAGGAGAGACGGTTGTCTCTGTAGAACTACCGTGTCTTCTGCTTGGCG	5760
Qy	5761	GAGCTTCGCACAAAGACCTTCGGCAGCTCCGAATCTGTCGGCGTCCGACGGCCACGGCA	5820
Db	5761	GAGCTTCGCACAAAGACCTTCGGCAGCTCCGAATCTGTCGGCGTCCGACGGCCACGGCA	5820
Qy	5821	ACGGCTCTCTGTACCACGCCCTCCGACGACGGCGACGCGGATCCGACGTGTAGTCGTAC	5880
Db	5821	ACGGCTCTCTGTACCACGCCCTCCGACGACGGCGACGCGGATCCGACGTGTAGTCGTAC	5880
Qy	5881	TCCTTCATGCCCCCTTGAGGGGGAGCGGGGGATCCCGATCTCAGCAGCGGTCTTGG	5940
Db	5881	TCCTTCATGCCCCCTTGAGGGGGAGCGGGGGATCCCGATCTCAGCAGCGGTCTTGG	5940
Qy	5941	TCTACCGTAAAGCAGGAGGCTAGTGAGGACGTCCTGCTGCTGCATGTCTTACACATGG	6000
Db	5941	TCTACCGTAAAGCAGGAGGCTAGTGAGGACGTCCTGCTGCTGCATGTCTTACACATGG	6000
Qy	6001	ACAGCGCCCTGATCAGCCATGCGCTGCGAGGAAACCAAGCTGCCATCTCATGCACTG	6060
Db	6001	ACAGCGCCCTGATCAGCCATGCGCTGCGAGGAAACCAAGCTGCCATCTCATGCACTG	6060
Qy	6061	AGCAACTCTTTGCTCCGTACCAACACTTGCTGTATGTCTACACATCTGCAGCGCAAGC	6120
Db	6061	AGCAACTCTTTGCTCCGTACCAACACTTGCTGTATGTCTACACATCTGCAGCGCAAGC	6120
Qy	6121	CTCGGCGAGGAAGGTCACTTTTGACAGACTGCAGGTCTTGACGACCACTACCGGAC	6180
Db	6121	CTCGGCGAGGAAGGTCACTTTTGACAGACTGCAGGTCTTGACGACCACTACCGGAC	6180
Qy	6181	GTGCTCAAGAGATGAAGGGAAGGGCTGCAAGTTAAGCTAACTTCTATCCGTGGAG	6240
Db	6181	GTGCTCAAGAGATGAAGGGAAGGGCTGCAAGTTAAGCTAACTTCTATCCGTGGAG	6240
Qy	6241	GAAGCTGTAAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGCTATGGGCAAG	6300
Db	6241	GAAGCTGTAAAGCTGACGCCCCACATTCGGCCAGATCTAAATTTGGCTATGGGCAAG	6300
Qy	6301	GAGTCOCGGAACCTATCCAGCAAGGCGTTAAACCAATCCGCTCCGTGTGGAAGACTTG	6360
Db	6301	GAGTCOCGGAACCTATCCAGCAAGGCGTTAAACCAATCCGCTCCGTGTGGAAGACTTG	6360
Qy	6361	CTGGAGACACTGAGACACCAATTGACACCAATCATGGCAAAATAGAGTTTCTGC	6420
Db	6361	CTGGAGACACTGAGACACCAATTGACACCAATCATGGCAAAATAGAGTTTCTGC	6420
Qy	6421	GTCCAAACGAGAGAGGGGGCGCAAGCCAGCTCCGCTTATCGTATCCAGATTTGGGG	6480
Db	6421	GTCCAAACGAGAGAGGGGGCGCAAGCCAGCTCCGCTTATCGTATCCAGATTTGGGG	6480
Qy	6481	GTTTCGTGTGCGAGAAATGGCCCTTTACGATGGGTCTCCACCTCCCTCAGGCCGTG	6540
Db	6481	GTTTCGTGTGCGAGAAATGGCCCTTTACGATGGGTCTCCACCTCCCTCAGGCCGTG	6540
Qy	6541	ATGGGCTCTTTACAGGATTCGAATCTCTCTGACAGCGGGTTCGAGTCTCTGCTGAAT	6600
Db	6541	ATGGGCTCTTTACAGGATTCGAATCTCTCTGACAGCGGGTTCGAGTCTCTGCTGAAT	6600
Qy	6601	GCTTGAAGCGAAGAAATGCCCCTATGGCTTCGCATATGACACCGCTGTTTTCATCA	6660
Db	6601	GCTTGAAGCGAAGAAATGCCCCTATGGCTTCGCATATGACACCGCTGTTTTCATCA	6660
Qy	6661	ACGGTCACTGAGAAATGACATCGTGTTCAGGAGTCAATCTACCAATGTTGTGACTTGCC	6720
Db	6661	ACGGTCACTGAGAAATGACATCGTGTTCAGGAGTCAATCTACCAATGTTGTGACTTGCC	6720
Qy	6721	CCGGAAGCCAGACAGCCCATAGTTCGTCTACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780
Db	6721	CCGGAAGCCAGACAGCCCATAGTTCGTCTACAGAGCGGCTTTACATCGGGGGCCCCCTG	6780

6781 ACTAATCTAAAGGACAGAACTGCGGCTATGCGCGTTCGCGGAGCGGTGTACTGACG 6840  
 Db  
 6781 ACTAATCTAAAGGACAGAACTGCGGCTATGCGCGTTCGCGGAGCGGTGTACTGACG 6840  
 Qy  
 6841 ACCAGCTGGGTAAATACCTTACATGTTACTTGAAGGCGGTGCGGCGGTGTGAGCTGG 6900  
 Db  
 6841 ACCAGCTGGGTAAATACCTTACATGTTACTTGAAGGCGGTGCGGCGGTGTGAGCTGG 6900  
 Qy  
 6901 AAGCTCCAGGACTGACAGGATGCTGCTATGCGGAGACGACTTGTGCTTATCTGTAAGC 6960  
 Db  
 6901 AAGCTCCAGGACTGACAGGATGCTGCTATGCGGAGACGACTTGTGCTTATCTGTAAGC 6960  
 Qy  
 6961 GCGGGGACCCAAAGAGAGAGGCGGCTTACGCGGCTTACGAGGCTAAGCTAGATAC 7020  
 Db  
 6961 GCGGGGACCCAAAGAGAGAGGCGGCTTACGCGGCTTACGAGGCTAAGCTAGATAC 7020  
 Qy  
 7021 TCTGCCCCCTGGGACCCGCGCCAAACAGATAGCACTTGGAGTTGATAATCATATGC 7080  
 Db  
 7021 TCTGCCCCCTGGGACCCGCGCCAAACAGATAGCACTTGGAGTTGATAATCATATGC 7080  
 Qy  
 7081 TCTTCAATGTGTCACTGCGGACGATGCTTGGCAAAAGGCTGTACTATCTCACCGT 7140  
 Db  
 7081 TCTTCAATGTGTCACTGCGGACGATGCTTGGCAAAAGGCTGTACTATCTCACCGT 7140  
 Qy  
 7141 GACCCACACCCCTTGGCGGGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
 Db  
 7141 GACCCACACCCCTTGGCGGGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
 Qy  
 7201 TCTGTCTAGGCAACATCATGATGATGCGCCACCTTGGGCAAGGATGATCCTGATG 7260  
 Db  
 7201 TCTGTCTAGGCAACATCATGATGATGCGCCACCTTGGGCAAGGATGATCCTGATG 7260  
 Qy  
 7261 ACTCATTTCTTCCATCTTCTAGCTCAGGAACTTGAAGGCTTGAATTTGTCAG 7320  
 Db  
 7261 ACTCATTTCTTCCATCTTCTAGCTCAGGAACTTGAAGGCTTGAATTTGTCAG 7320  
 Qy  
 7321 ATCTAGGGGCTGTACTCCATTTAGGACCACTTGAACCTCAGATCATCAACACTC 7380  
 Db  
 7321 ATCTAGGGGCTGTACTCCATTTAGGACCACTTGAACCTCAGATCATCAACACTC 7380  
 Qy  
 7381 CATGGCTTAGCGATTTTCACTCCATAGTACTTCCAGGTGAGATCAATAGGTGGT 7440  
 Db  
 7381 CATGGCTTAGCGATTTTCACTCCATAGTACTTCCAGGTGAGATCAATAGGTGGT 7440  
 Qy  
 7441 TCATGCTCAGGAACTTGGGCTACCGCTTGGAGTCTGGAGACATCGGGCCAGAGT 7500  
 Db  
 7441 TCATGCTCAGGAACTTGGGCTACCGCTTGGAGTCTGGAGACATCGGGCCAGAGT 7500  
 Qy  
 7501 GTCGCGCTAGGCTTACTGTCAGGGGGAGGCTGCCCTTGTGGCAAGTACTCTTC 7560  
 Db  
 7501 GTCGCGCTAGGCTTACTGTCAGGGGGAGGCTGCCCTTGTGGCAAGTACTCTTC 7560  
 Qy  
 7561 AACTGGGAGTAAAGCAAGCTCAAACTCACTCCATCCCGGCTGGCTCCAGTTCGAT 7620  
 Db  
 7561 AACTGGGAGTAAAGCAAGCTCAAACTCACTCCATCCCGGCTGGCTCCAGTTCGAT 7620  
 Qy  
 7621 TTATCCAGCTGGTTCGTTGCTGTGTTACAGGGGGAGACATATACAGCTGTCTCGT 7680  
 Db  
 7621 TTATCCAGCTGGTTCGTTGCTGTGTTACAGGGGGAGACATATACAGCTGTCTCGT 7680  
 Qy  
 7681 GCCCGACCCGCTGGTTCATGTGTGCTACTCTTCTGTTAGGGGTAGGATCATAT 7740  
 Db  
 7681 GCCCGACCCGCTGGTTCATGTGTGCTACTCTTCTGTTAGGGGTAGGATCATAT 7740  
 Qy  
 7741 CTACTCCCAACCGATGAAGGAGCTTAAACACTCCAGGCAATAGGCCATCTGTTT 7800  
 Db  
 7741 CTACTCCCAACCGATGAAGGAGCTTAAACACTCCAGGCAATAGGCCATCTGTTT 7800  
 Qy  
 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
 Db  
 7801 TTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860

7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
 Db  
 7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
 Qy  
 7921 TAGCTGTGAAAGGTCGTTGAGCGGCTTACCTGACAGAGTGTGATCTGGCTTCTTCG 7980  
 Db  
 7921 TAGCTGTGAAAGGTCGTTGAGCGGCTTACCTGACAGAGTGTGATCTGGCTTCTTCG 7980  
 Qy  
 7981 AGATCAAGT 7989  
 Db  
 7981 AGATCAAGT 7989

RESULT 10  
 US-10-005-469-5  
 ; Sequence 5, Application US/10005469  
 ; Publication No. US20020155133A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: ANADYS Pharmaceuticals, Inc.  
 ; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REP  
 ; FILE REFERENCE: 0342/1H395US1  
 ; CURRENT APPLICATION NUMBER: US/10/005,469  
 ; CURRENT FILING DATE: 2002-04-18  
 ; PRIOR APPLICATION NUMBER: US 60/245,866  
 ; PRIOR FILING DATE: 2000-11-07  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: Patent in version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 7992  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HCV Replicon from cell line HCVR22  
 US-10-005-469-5

Query Match 99.9%; Score 7982.6; DB 13; Length 7992;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 7985; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 Db 1 GCCAGCCCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60  
 Qy 61 TCTTCAGCAGAAAGCGTCTAGCCATGGCTTGTAGTGTGCTGTCAGCTCCAGGAC 120  
 Db 61 TCTTCAGCAGAAAGCGTCTAGCCATGGCTTGTAGTGTGCTGTCAGCTCCAGGAC 120  
 Qy 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180  
 Db 121 CCCCCCTCCCGGAGAGCCATAGTGTCTCGGAAACCGGTGAGTACACCGGAATTGCCAG 180  
 Qy 181 GAGCAGCGGTCTTTCTTGATCAACCCCTCAATGCTGGAGATTGGCGGTGCCCCC 240  
 Db 181 GAGCAGCGGTCTTTCTTGATCAACCCCTCAATGCTGGAGATTGGCGGTGCCCCC 240  
 Qy 241 GCGAGACTGTAGCCGAGTGTGGTTCGAAAGCGCTTGTGTACTGCCCTGATAGG 300  
 Db 241 GCGAGACTGTAGCCGAGTGTGGTTCGAAAGCGCTTGTGTACTGCCCTGATAGG 300  
 Qy 301 GTGCTTCGAGTGTCCCGGAGGCTCTGTAGACCGGTGACCATGAGCAGCAATCCTAAAC 360  
 Db 301 GTGCTTCGAGTGTCCCGGAGGCTCTGTAGACCGGTGACCATGAGCAGCAATCCTAAAC 360  
 Qy 361 CTCAGAGAAAACCAAGGGCGGCCATGATTGAACAGATGATTGACGACGCTCTC 420  
 Db 361 CTCAGAGAAAACCAAGGGCGGCCATGATTGAACAGATGATTGACGACGCTCTC 420  
 Qy 421 CGGCGCTTGGGTGAGAGGCTATTTCGCTATGATGGGCAACAGCAATTCGCTGCT 480  
 Db 421 CGGCGCTTGGGTGAGAGGCTATTTCGCTATGATGGGCAACAGCAATTCGCTGCT 480  
 Qy 481 CTGATGCGCGCTGTTCGGGCTGTACGCGAGGGGGCGGCTCTTTTGTCAAGCGG 540



Db	481	CTGATGCGCGCGTGTTCGGGCTGTACGCGCAGGGGCGCCGGGTTCTTTTGTTCGAAGACCG	540
Qy	541	ACCTGTCCGGTGCCTGAATGAACTGACAGCAGAGCAGCGCGGTATCTGTGGCTTGGCCA	600
Db	541	ACCTGTCCGGTGCCTGAATGAACTGACAGCAGAGCAGCGCGGTATCTGTGGCTTGGCCA	600
Qy	601	CGACGGGCGTTCCTTTCGCGACGTGTCTCGACGTGTCTCACTGAACGGGAAAGGACTGGC	660
Db	601	CGACGGGCGTTCCTTTCGCGACGTGTCTCGACGTGTCTCACTGAACGGGAAAGGACTGGC	660
Qy	661	TGCTATTGGGCGAAGTCCGCGGGCAGAGTCTCCTGTCTATCTCACTTGTCTGTCCGCGAGA	720
Db	661	TGCTATTGGGCGAAGTCCGCGGGCAGAGTCTCCTGTCTATCTCACTTGTCTGTCCGCGAGA	720
Qy	721	AAGTATCCATCATGGCTGATGCAATTCGGCGCGCTGCATACGCTTGTATCCGGCTACCTGCC	780
Db	721	AAGTATCCATCATGGCTGATGCAATTCGGCGCGCTGCATACGCTTGTATCCGGCTACCTGCC	780
Qy	781	CATTTCGACCAAGCGGAAACATCGCATTCGACGAGCACGTACTCGATGGAAGCCGGTC	840
Db	781	CATTTCGACCAAGCGGAAACATCGCATTCGACGAGCACGTACTCGATGGAAGCCGGTC	840
Qy	841	TTGTTCGATCAGGATGATCTGACCAAGAGCATCAGGGGCTTCGGCCAGCGCAACTGTTCG	900
Db	841	TTGTTCGATCAGGATGATCTGACCAAGAGCATCAGGGGCTTCGGCCAGCGCAACTGTTCG	900
Qy	901	CCAGGCTCAAAGCGCGCATGCCCGACGCGGAGGATCTCGTCGTGACCCATGCGCATGCGCT	960
Db	901	CCAGGCTCAAAGCGCGCATGCCCGACGCGGAGGATCTCGTCGTGACCCATGCGCATGCGCT	960
Qy	961	GCTTCCCGAATATCATGGTGGAAATAGCCCGCTTTCTTGGATTCATCGACTGTGCCCGCG	1020
Db	961	GCTTCCCGAATATCATGGTGGAAATAGCCCGCTTTCTTGGATTCATCGACTGTGCCCGCG	1020
Qy	1021	TGGGTGTGGCGGACCGCTATCAGACATACGTTGGCTACCCGTGATATGCTGAAGAGC	1080
Db	1021	TGGGTGTGGCGGACCGCTATCAGACATACGTTGGCTACCCGTGATATGCTGAAGAGC	1080
Qy	1081	TTGGCGCGCAATGGGCTGACCGCTTCCTCGTGTCTTACGGTATCGCCGCTCCCGATTTCG	1140
Db	1081	TTGGCGCGCAATGGGCTGACCGCTTCCTCGTGTCTTACGGTATCGCCGCTCCCGATTTCG	1140
Qy	1141	AGCGATTCGCTTCATTCGCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACAAG	1200
Db	1141	AGCGATTCGCTTCATTCGCTTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACAAG	1200
Qy	1201	GTTTTCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCTTAAGTACTGGC	1260
Db	1201	GTTTTCCTCTAGCGGGATCAATTCGCGCCCTCTCCCTCCCGCCCGCTTAAGTACTGGC	1260
Qy	1261	CGAAGCGCTTGGAAATAAGCCCGGTGTGCTTGTCTATATGTATTTTCCACCATATTG	1320
Db	1261	CGAAGCGCTTGGAAATAAGCCCGGTGTGCTTGTCTATATGTATTTTCCACCATATTG	1320
Qy	1321	CCGTCTTTTGGCAATGTAGGGCCCGGAAACCTTGCGCTCTCTTCTTGACGAGCATTCCT	1380
Db	1321	CCGTCTTTTGGCAATGTAGGGCCCGGAAACCTTGCGCTCTCTTCTTGACGAGCATTCCT	1380
Qy	1381	AGGGGTCTTTCCCTCTCGCAAGGATTCAGGTCCTGTGATGTCTGTGAGAGGAGCA	1440
Db	1381	AGGGGTCTTTCCCTCTCGCAAGGATTCAGGTCCTGTGATGTCTGTGAGAGGAGCA	1440
Qy	1441	GTTTCTCTGTGAAGCTTCTTGAAGACAAACACGTCTGACGACCCCTTTCGACGACGCG	1500
Db	1441	GTTTCTCTGTGAAGCTTCTTGAAGACAAACACGTCTGACGACCCCTTTCGACGACGCG	1500
Qy	1501	AACCCCCCACTGGGACAGGTGCCTCTCGGCGCAAAAGCCAGCTGTATTAAGATACACCT	1560
Db	1501	AACCCCCCACTGGGACAGGTGCCTCTCGGCGCAAAAGCCAGCTGTATTAAGATACACCT	1560
Qy	1561	GCAAAGGGCGCAACCCAGTCCACGTTGTGAGTTGATAGTTGGAAAGAGTCAAA	1620

Db	1561	GCAAGGCGCACAAACCAGTCGCACGTTGTGAGTTGGATAGTTGTGGAAAGAGTCAA	1620
Qy	1621	TGGCTCTCCTCAAGGTATTCAAAGGGGCTGAAGATGCCAGAAAGGTACCCCATTTGT	1680
Db	1621		1680
Db	1621	TGGCTCTCCTCAAGCGTATTCAAAGGGGCTGAAGATGCCAGAAAGGTACCCCATTTGT	1680
Qy	1681	ATGGGATCTGATCTCTGGGCGCTCGGTGCACATCTCTTTACATGTGTTTAACTCGAGGTTAAAA	1740
Db	1681		1740
Db	1681	ATGGGATCTGATCTGGGGCTCTCGTGACATGCTTTACATGTGTTTAACTCGAGGTTAAAA	1740
Qy	1741	AACGTCTAAGGCCCCCGAACAACGGGACGTGGTTTTCTTTGAAAAACACGATAATACC	1800
Db	1741	AACGTCTAAGGCCCCCGAATCACGGGACGTGGTTTTCTTTGAAAAACACGATAATACC	1800
Qy	1801	ATGGGCGCTATTACGGGCTACTCCCAACAGACGCGAGCGCTACTTTGCTGCATCATCACT	1860
Db	1801		1860
Db	1801	ATGGCGCTATTACGGGCTACTCCCAACAGACGCGAGCGCTACTTTGCTGCATCATCACT	1860
Qy	1861	AGCCTCACAGGCGGGGACAGAAACAGCTCGAGGGGAGGTCCAAGTGTCTCCACCGCA	1920
Db	1861	AGCCTCACAGGCGGGACAGAACACAGTTCGAGGGGAGGTCCAAGTGTCTCCACCGCA	1920
Qy	1921	ACACAACTTTCTCTGGGCAACTGGGTCAATGGCGGTGTGGACTGTCTATCATGTGCC	1980
Db	1921	ACAAATCTTTCTCTGGGCACTGGTCAATGGCGGTGTGGACTGTCTATCATGTGCC	1980
Qy	1981	GGCTCAAGAGCCCTTCGGCGCCCAAGGGCGCGTTCCTTGACACCATGCACCTGC	2040
Db	1981	GGCTCAAGAGCCCTTCGGCGCCCAAGGGCGCGTTCCTTGACACCATGCACCTGC	2040
Qy	2041	CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGCGCGTTCCTTGACACCATGCACCTGC	2100
Db	2041	CAGGACCTCGTGGCTGGCAAGCGCCCCCGGGCGCGTTCCTTGACACCATGCACCTGC	2100
Qy	2101	GGCAGCTCGGACCTTTACTTGGTCAAGAGCATGCCGATGTCATTCGGGTGCGCGCGG	2160
Db	2101	GSCAGCTCGGACCTTTACTTGGTCAAGAGCATGCCGATGTCATTCGGGTGCGCGCGG	2160
Qy	2161	GGCCACAGCAGGGGGAGCTACTCTCCCCAGCGCCGTCTCTTCTTGAAGGCTCTTGG	2220
Db	2161	GGCCACAGCAGGGGGAGCCTACTCTCCCCAGCGCCGTCTCTTCTTGAAGGCTCTTGG	2220
Qy	2221	GGCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGATCTTTCGGGCTGCCGCTGC	2280
Db	2221	GGCGGTCCACTGCTCTGCCCTCGGGGACGCTGTGGGATCTTTCGGGCTGCCGCTGC	2280
Qy	2281	ACCCGAGGGGTGCGAAGGCGGTGGACTTTGTACCCGTCGAGTCTATGGAAACCAATG	2340
Db	2281	ACCCGAGGGGTGCGAAGGCGGTGGACTTTGTACCCGTCGAGTCTATGGAAACCAATG	2340
Qy	2341	CGFTCCCGGCTTTCAGGACACTCGTCCCCCGCGGTACCGGACAGATTTCCAGGTG	2400
Db	2341	CGFTCCCGGCTTTCAGGACAACTCGTCCCCCGCGGTACCGGACAGATTTCCAGGTG	2400
Qy	2401	GCCCATCTACGCCCTCTACTGTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA	2460
Db	2401	GCCCATCTACGCCCTCTACTGTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA	2460
Qy	2461	GCCCAAGGGTAAAGTGTGTGCTCTGAAACCGTCCGTCGCGCCACCTAGGTTTCGGG	2520
Db	2461	GCCCAAGGGTAAAGTGTGTGCTCTGAAACCGTCCGTCGCGCCACCTAGGTTTCGGG	2520
Qy	2521	GCCTATATGCTTAAGGCACATGTTATCGACCTTAACATCAGAACCGGGTAAGGACCATC	2580
Db	2521	GCCTATATGCTTAAGGCACATGTTATCGACCTTAACATCAGAACCGGGTAAGGACCATC	2580
Qy	2581	ACCACGGGTGCCCCCATCAGTACTCCACTATGGCAAGTTTCTTTCGCGACGCTGTTGC	2640
Db	2581	ACCACGGGTGCCCCCATCAGTACTCCACTATGGCAAGTTTCTTTCGCGACGCTGTTGC	2640
Qy	2641	TTCTGGGGCGCCTATGACATCATATATGTATGTAGTGCCACTCAACTGACTCCACCACT	2700
Db	2641	TTCTGGGGCGCCTATGACATCATATATGTATGTAGTGCCACTCAACTGACTCCACCACT	2700



Db 4861 ATCTGGGGCCCTAGGACCTGTAGTAACACGTCGCAATGGAACATCTCCCAATTAACGGCTAC 4920  
QY 4921 ACCACGGCCCTGACAGCCCTCCCGGGCCCAATATTCTAGGGCGCTGTGGCGGGTG 4980  
Db 4921 ACCACGGCCCTGACAGCCCTCCCGGGCCCAATATTCTAGGGCGCTGTGGCGGGTG 4980  
QY 4981 GCTGCTGAGGAGTACGTCGAGGTTACGGGGGTGGGGGATTTCCATACGTGACGGGCATG 5040  
Db 4981 GCTGCTGAGGAGTACGTCGAGGTTACGGGGGTGGGGGATTTCCATACGTGACGGGCATG 5040  
QY 5041 ACCACTGACAGGTAAGTGCCCGTGTGAGGTTCCGGGCCCGCAATTTCTTACAGAAGTG 5100  
Db 5041 ACCACTGACAGGTAAGTGCCCGTGTGAGGTTCCGGGCCCGCAATTTCTTACAGAAGTG 5100  
QY 5101 GATGGGGTGGGGTGGCAGAGTACGTCAGGTTCCAGGTCGAACCCCTCTACGGGAGGAGTGC 5160  
Db 5101 GATGGGGTGGGGTGGCAGAGTACGTCAGGTTCCAGGTCGAACCCCTCTACGGGAGGAGTGC 5160  
QY 5161 ACATTCCTGGTCCGGCTCAATCAATACCTGTTGGGTCCACAGCTCCCATGCGAGCCCGAA 5220  
Db 5161 ACATTCCTGGTCCGGCTCAATCAATACCTGTTGGGTCCACAGCTCCCATGCGAGCCCGAA 5220  
QY 5221 CCGAGAGTAGAGTCTCACTTCCATGCTCACCGACCCCTCCCAATTAACGGGCAGG 5280  
Db 5221 CCGAGAGTAGAGTCTCACTTCCATGCTCACCGACCCCTCCCAATTAACGGGCAGG 5280  
QY 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTGGCCAGCTCATAGCTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTGGCCAGCTCATAGCTAGCCAG 5340  
QY 5341 CTGTCTGGCCCTCTTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGCGCTGAC 5400  
Db 5341 CTGTCTGGCCCTCTTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGCGCTGAC 5400  
QY 5401 CTGATCAGAGGCAACCTCTGTCGGCAGAGATGGCGGGAAACATCACCCCGGTGGAG 5460  
Db 5401 CTGATCAGAGGCAACCTCTGTCGGCAGAGATGGCGGGAAACATCACCCCGGTGGAG 5460  
QY 5461 TCAGAAATAGGTAGTAATTTTGGACTCTTTTCAGCGCTTCCAGCGGAGGAGATGAG 5520  
Db 5461 TCAGAAATAGGTAGTAATTTTGGACTCTTTTCAGCGCTTCCAGCGGAGGAGATGAG 5520  
QY 5521 AGGAAAGTATCCGTTCCGGCGGAGATCTCGGGAGGTCAGGAAATTTCCCTCGAGCGATG 5580  
Db 5521 AGGAAAGTATCCGTTCCGGCGGAGATCTCGGGAGGTCAGGAAATTTCCCTCGAGCGATG 5580  
QY 5581 CCCATATGGGACGCCCGGATTAACAACCTTCACTGTTAGAGTCTTGGAAAGGACCCGGAC 5640  
Db 5581 CCCATATGGGACGCCCGGATTAACAACCTTCACTGTTAGAGTCTTGGAAAGGACCCGGAC 5640  
QY 5641 TAGCTCCCTCCAGTGTACAGGGTGTCCATGCGGCTCGCCAAAGGCCCTCCGATACCA 5700  
Db 5641 TAGCTCCCTCCAGTGTACAGGGTGTCCATGCGGCTCGCCAAAGGCCCTCCGATACCA 5700  
QY 5701 CCTCCAGGAGGAGGACGTTGTCTCTGTCAGAACTACCGGTGTCTTCTGCTGGCG 5760  
Db 5701 CCTCCAGGAGGAGGACGTTGTCTCTGTCAGAACTACCGGTGTCTTCTGCTGGCG 5760  
QY 5761 GAGCTGCCACAAAGACCTTCGACAGTCCGAAATCGTCGGCCGTTCGACAGCGGACGGCA 5820  
Db 5761 GAGCTGCCACAAAGACCTTCGACAGTCCGAAATCGTCGGCCGTTCGACAGCGGACGGCA 5820  
QY 5821 ACGGCTCTCTGACAGCCCTCGACAGCGGAGTCCGGGATCCGACGTTGAGTGTGATC 5880  
Db 5821 ACGGCTCTCTGACAGCCCTCGACAGCGGAGTCCGGGATCCGACGTTGAGTGTGATC 5880  
QY 5881 TCCTCATGCCCCCTTTAGGGGGAGCCGGGGATCCCGATCTCAGCGACGGGTCTTGG 5940  
Db 5881 TCCTCATGCCCCCTTTAGGGGGAGCCGGGGATCCCGATCTCAGCGACGGGTCTTGG 5940  
QY 5941 TCTACCGTAAGCGAGGAGGCTAGTGAGGAGCTGCTGTGCTGTGCTCGATGTCTTACATAGG 6000

Db 5941 TCTACCGTAAGCGAGGAGGCTAGTGAGGAGCTGCTGCTGCTCGATGTCTTACATGG 6000  
QY 6001 ACAGCGCCCTGATCACGCCATCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
Db 6001 ACAGCGCCCTGATCACGCCATCGCTCGGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
QY 6061 AGCAACTCTTTGCTCCGTCACCAACACTTGTCTATGCTACAACTCTCGAGCGCAAGC 6120  
Db 6061 AGCAACTCTTTGCTCCGTCACCAACACTTGTCTATGCTACAACTCTCGAGCGCAAGC 6120  
QY 6121 CTGGCGCAAGAGAGGTCACCTTTGACAGACTGACAGTCTCGGAGACCACTACCGGAC 6180  
Db 6121 CTGGCGCAAGAGAGGTCACCTTTGACAGACTGACAGTCTCGGAGACCACTACCGGAC 6180  
QY 6181 GTGCTCAAGAGAGTGAAGCGGAAGCGCTCCACAGTTAAGGCTAAACTTCTATCCGTGGAG 6240  
Db 6181 GTGCTCAAGAGAGTGAAGCGGAAGCGCTCCACAGTTAAGGCTAAACTTCTATCCGTGGAG 6240  
QY 6241 GAAGCTGTAAAGCTGACGCCGCCACATTCGGCCAGATCTAAATTTGGCTATGGGCCAAG 6300  
Db 6241 GAAGCTGTAAAGCTGACGCCGCCACATTCGGCCAGATCTAAATTTGGCTATGGGCCAAG 6300  
QY 6301 GACGTCCGGAACCTTATCCAGCAAGCGCTTAAACACATCCGCTCGGTGGAAGGACTTG 6360  
Db 6301 GACGTCCGGAACCTTATCCAGCAAGCGCTTAAACACATCCGCTCGGTGGAAGGACTTG 6360  
QY 6361 CTGGAAGACACTGAGACACCAATTTGACACCACTCATGTCGCAAAATGAGTTTCTGC 6420  
Db 6361 CTGGAAGACACTGAGACACCAATTTGACACCACTCATGTCGCAAAATGAGTTTCTGC 6420  
QY 6421 GTCCAAACAGAGAGGGGGCGGCAAGCGCTCGCTTATGCTATTTCCAGATTTGGGG 6480  
Db 6421 GTCCAAACAGAGAGGGGGCGGCAAGCGCTCGCTTATGCTATTTCCAGATTTGGGG 6480  
QY 6481 GTTCGTGTGTCGAGAAATGGCCCTTTACAGTGTGCTTCCACCTTCCCTCAGGCGGTG 6540  
Db 6481 GTTCGTGTGTCGAGAAATGGCCCTTTACAGTGTGCTTCCACCTTCCCTCAGGCGGTG 6540  
QY 6541 ATGGGCTCTTTCATACGGAATTCGAATCTCTCTGGACAGCGGGTCGAGTTCCTGGTGAAT 6600  
Db 6541 ATGGGCTCTTTCATACGGAATTCGAATCTCTCTGGACAGCGGGTCGAGTTCCTGGTGAAT 6600  
QY 6601 GCCTGGAAGCGGAAGAAATGCCCTATGCGGTTCGCAATGACACCCGCTGTTTGACTCA 6660  
Db 6601 GCCTGGAAGCGGAAGAAATGCCCTATGCGGTTCGCAATGACACCCGCTGTTTGACTCA 6660  
QY 6661 ACGGTCACTCAGAAATGACATCCGTTGTTGAGGAGTCAATCTACCAATGTTGTGACTGGCC 6720  
Db 6661 ACGGTCACTCAGAAATGACATCCGTTGTTGAGGAGTCAATCTACCAATGTTGTGACTGGCC 6720  
QY 6721 CCCGAAGCCAGACAGGCCATTAAGTTCGCTCACAGAGCGGCTTACATCGGGGGCCCTG 6780  
Db 6721 CCCGAAGCCAGACAGGCCATTAAGTTCGCTCACAGAGCGGCTTACATCGGGGGCCCTG 6780  
QY 6781 ACTAATTTAAAGGGCAGAACTCGCGGTATCGCGGTGCGCGAGCGGTGTACTGAACG 6840  
Db 6781 ACTAATTTAAAGGGCAGAACTCGCGGTATCGCGGTGCGCGAGCGGTGTACTGAACG 6840  
QY 6841 ACCAGCTGCGGTAAATACCCCTCACATGTTACTTGAAGCGCGCTCGGGCTGTGAGCTGGC 6900  
Db 6841 ACCAGCTGCGGTAAATACCCCTCACATGTTACTTGAAGCGCGCTCGGGCTGTGAGCTGGC 6900  
QY 6901 AAGCTCCAGGACTGACGATGCTGCTGATCGGAGAGCACTTCTGCTTATCTGTGAAGC 6960  
Db 6901 AAGCTCCAGGACTGACGATGCTGCTGATCGGAGAGCACTTCTGCTTATCTGTGAAGC 6960  
QY 6961 GCGGGACCCAGAGAGAGCGAGCGCTTACCGGCTTCCAGGAGGCTATGACTAGATAC 7020  
Db 6961 GCGGGACCCAGAGAGAGCGAGCGCTTACCGGCTTCCAGGAGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTTGGGAGCCGCCCAACAGAACTAGACTTGGAGTTGATAATCATGTC 7080  
Db 7021 TCTGCCCCCTTGGGAGCCGCCCAACAGAACTAGACTTGGAGTTGATAATCATGTC 7080

```

7081 TCCTCCAAATGTGTCAGTCCGCGACGATGCATCTGGCAAAGGGTGTACTATCTCACCCGT 7140
7081 TCCTCCAAATGTGTCAGTCCGCGACGATGCATCTGGCAAAGGGTGTACTATCTCACCCGT 7140
7141 GACCCACACACCCCTTTCGCGGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
7141 GACCCACACACCCCTTTCGCGGGCTGCTGGGAGACAGCTAGACACACTCCAGTCAAT 7200
7201 TCCTGGCTAGGCAACATCATCATATGATGCGCCCACTTGTGGCAAGGATGATCTCTGATG 7260
7201 TCCTGGCTAGGCAACATCATCATATGATGCGCCCACTTGTGGCAAGGATGATCTCTGATG 7260
7261 ACTCAATTTCTTCATCTTCTAGCTCAGGACAACTTGAAGAGCCCTAGATGTCAG 7320
7261 ACTCAATTTCTTCATCTTCTAGCTCAGGACAACTTGAAGAGCCCTAGATGTCAG 7320
7321 ATCTAGGGGCTGTGTACTCCATTGAGCCACTTGAACCTACCTCAGATCAATCAAGACTC 7380
7321 ATCTAGGGGCTGTGTACTCCATTGAGCCACTTGAACCTACCTCAGATCAATCAAGACTC 7380
7381 CATGGCTTAGGCGATTTTCACTCCATAGTTACTTCCAGGTGAGATCAATAGGGTGGCT 7440
7381 CATGGCTTAGGCGATTTTCACTCCATAGTTACTTCCAGGTGAGATCAATAGGGTGGCT 7440
7441 TCATGCTCAGGAACTTGGGGTACCGCCCTTGCAGTCTGGAGACATCGGGCCAGAGT 7500
7441 TCATGCTCAGGAACTTGGGGTACCGCCCTTGCAGTCTGGAGACATCGGGCCAGAGT 7500
7501 GTCCGGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560
7501 GTCCGGCTAGGCTACTGTCCAGGGGGAGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560
7561 AACTGGGAGTAGAGCAAGCTCAAACTCACTCCAAATCCCGGCTGCTGCCAGTTGGAT 7620
7561 AACTGGGAGTAGAGCAAGCTCAAACTCACTCCAAATCCCGGCTGCTGCCAGTTGGAT 7620
7621 TTATCCAGCTGGTTCGTGTGCTGTACAGGGGGAGACATATACAGCCCTGTCTCGT 7680
7621 TTATCCAGCTGGTTCGTGTGCTGTACAGGGGGAGACATATACAGCCCTGTCTCGT 7680
7681 GCGCGACCCGCTGTGTGTGCTGTACTCTACTTCTGTAGGGGTAGGCATCTAT 7740
7681 GCGCGACCCGCTGTGTGTGCTGTACTCTACTTCTGTAGGGGTAGGCATCTAT 7740
7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800
7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800
7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860
7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
7861 TTTTCTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920
7921 TAGCTGTGAAGGTCCTGTAGCCGCTTGAATGAGAGAGTGTGATAGTGGCTCTCTGC 7980
7921 TAGCTGTGAAGGTCCTGTAGCCGCTTGAATGAGAGAGTGTGATAGTGGCTCTCTGC 7980
7981 AGATCAAGT 7989
7981 AGATCAAGT 7989

```

RESULT 11  
 US-10-434-842-2  
 ; Sequence 2, Application US/10434842  
 ; Publication No. US2004005549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REH

```

; FILE REFERENCE: 0342/1H395US3
; CURRENT APPLICATION NUMBER: US/10/434,842
; CURRENT FILING DATE: 2003-05-09
; PRIOR APPLICATION NUMBER: US 10/233,307
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: US 10/005,469
; PRIOR FILING DATE: 2001-11-07
; PRIOR APPLICATION NUMBER: US 60/245,866
; PRIOR FILING DATE: 2000-11-07
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 7992
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: HCV2 subgenomic HCV replicon
; US-10-434-842-2

Query Match 99.9%; Score 7982.6; DB 16; Length 7992;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 7985; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GCCAGCCCCGATTTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60
DB 1 GCCAGCCCCGATTTGGGGGCGACACTCCACCATAGATCACTCCCTGTGAGGAACACTG 60
QY 61 TCTTACGCGAGAAAGGCTTAGCCATGCGGTAGTATGATGTCGTCAGCCCTCCAGGAC 120
DB 61 TCTTACGCGAGAAAGGCTTAGCCATGCGGTAGTATGATGTCGTCAGCCCTCCAGGAC 120
QY 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCG 180
DB 121 CCCCCCTCCCGGAGAGCCATAGTGTCTGCGAAACCGGTGAGTACACCGGAATTCGCG 180
QY 181 GAGACCGGGTCTTTTCTTGGATCAACCCGCTCAATGCCCTGAGATTTGGGCGTGCCTCC 240
DB 181 GAGACCGGGTCTTTTCTTGGATCAACCCGCTCAATGCCCTGAGATTTGGGCGTGCCTCC 240
QY 241 GCGAGACTCTAGCCGAGTAGTGTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300
DB 241 GCGAGACTCTAGCCGAGTAGTGTGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 300
QY 301 GTGCTTTCGAGTGCCTCCGCGAGGTCTCGTAGACCGTGCACCATGAGCAGAACTCTAAAC 360
DB 301 GTGCTTTCGAGTGCCTCCGCGAGGTCTCGTAGACCGTGCACCATGAGCAGAACTCTAAAC 360
QY 361 CTCAAGAAAAAACAAGGGCGGCCATGATTTGAACAAGATGATTTGACGCGAGGTCTC 420
DB 361 CTCAAGAAAAAACAAGGGCGGCCATGATTTGAACAAGATGATTTGACGCGAGGTCTC 420
QY 421 CGGCGCTTGGTGGAGAGGCTATTGCGCTATGCTGGGCACACAGACAAATCGGCTGCT 480
DB 421 CGGCGCTTGGTGGAGAGGCTATTGCGCTATGCTGGGCACACAGACAAATCGGCTGCT 480
QY 481 CTGATCGCGCGCTTTCGCGCTCTCAGCGCAGGGCGCGCGGTTCTTTTGTCAAGACCG 540
DB 481 CTGATCGCGCGCTTTCGCGCTCTCAGCGCAGGGCGCGCGGTTCTTTTGTCAAGACCG 540
QY 541 ACTGTTCGGTGCCTGATGAATGAATGAGACAGGAGCGCGGCTATCTGCTGCTGCTG 600
DB 541 ACTGTTCGGTGCCTGATGAATGAATGAGACAGGAGCGCGGCTATCTGCTGCTGCTG 600
QY 601 CGACGGCGTTCCTTTCGCGAGCTGTGCTCGACCTGTGCTGCTGCTGCTGCTGCTGCTG 660
DB 601 CGACGGCGTTCCTTTCGCGAGCTGTGCTCGACCTGTGCTGCTGCTGCTGCTGCTGCTG 660
QY 661 TGCTATTGGCGGAAGTGCAGGAGTCTCTGTCTCATCTCACTTGTCTCTGCTGCTGCTG 720
DB 661 TGCTATTGGCGGAAGTGCAGGAGTCTCTGTCTCATCTCACTTGTCTCTGCTGCTGCTG 720
QY 721 AAGTATCCATGCTGCTGATGCAATGCGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 780
DB 721 AAGTATCCATGCTGCTGATGCAATGCGGGCTGATGCTGCTGCTGCTGCTGCTGCTGCTG 780

```

Db	721	AAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATCGGCTACTCGC	780
Qy	781	CATTGACACCAAGCAACATCCATCGAGGACGAGTACTCGATGGAAGCGGTC	840
Db	781		
Qy	841	TTGTCGATCAGGATGATCTGACGAAGAGCATCAGGGGCTCGCGCAGCGCAACTGTTTCG	900
Db	841		
Qy	901	CCAGCTCAAGCGCGCATGCCCAGCGGAGGATCTCGTCTGTGACCCCATGCGCATGCT	960
Db	901		
Qy	961	GCTTCCGGAATATCATGTGGAAATGCGCGCTTTCTGGATTATCGACTGTGCGCGC	1020
Db	961		
Qy	1021	TGGGTGTGGCGACCGCTATCAGGACATAGCGTTGGCTACCGTGATATTGCTGAAGAGC	1080
Db	1021		
Qy	1081	TTGGCGGGAATGGCTGACCGCTTCTCGTCTTTACGATACCGCTCCCGATTGCG	1140
Db	1081		
Qy	1141	AGCGCATCGCTTCTATCGCTTCTTGACGAGTCTTCTGAGTTTAAACAGACACAAAG	1200
Db	1141		
Qy	1201	GTTTCCCTCTAGCGGATCAATTCCGCGCTCTCCCTCCCGCCCTTAAGCTTACTGCG	1260
Db	1201		
Qy	1261	CGAAGCGCTTGGAAATAGCGCGGTGCGTGTGCTATATGTTATTTCCACCATATTG	1320
Db	1261		
Qy	1321	CCGTCTTTTGGCAATGTGAGGCGCGGAACTGGCCCTGCTTCTTGACGAGCATTCCT	1380
Db	1321		
Qy	1381	AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGTCTGTTGAATGCTGAAGGAAGCA	1440
Db	1381		
Qy	1441	GTTCTCTGGAAGCTTCTTGAAGCAAAACAACTGTGTAGCGACCTTTTGACGAGCGG	1500
Db	1441		
Qy	1501	AACCCCGCATCTGGCGACAGTGCCTCTGGGCCAAAGCCACGCTGATAGATACACCT	1560
Db	1501		
Qy	1561	GCAAGCGGCACAAACCCAGTGCACGTTGTGATTTGGATGTTGGAAAGAGTCAAA	1620
Db	1561		
Qy	1621	TGGCTCTCAAGCTATTCAACAGGGGTGAAGATGCCAGAAAGGTACCCCATTTG	1680
Db	1621		
Qy	1681	ATGGATCTGATCTGGGGCTCGGTGCACATGCTTTACATGTTTGTAGTGAAGTTAAAA	1740
Db	1681		
Qy	1741	AAGCTTAGGCCCGCCGAAACCAAGGACGTGGTTTCTTTGAAACACAGATAATACC	1800
Db	1741		
Qy	1801	ATGGCGCTATTACGGCTTACTCCCAACAGACGAGGCGCTACTTGGCTGCATCATCT	1860
Db	1801		
Qy	1861	AGCCTCAAGGCGCGGACAGGAACAGGTGCGAGGGAGGTCAAAGTGTCTCCACCGCA	1920
Db	1861		
Qy	1921	ACACAATCTTTCCTGGCGACCTGCTCAATGGCGTGTGTGGACTGTCTATCATGTGCG	1980
Db	1921		
Qy	1981	GGCTCAAAAGACCTTTGCGCGGCCCAAAAGGCGCAATCAACCAAAATGACCAAAATGTGAC	2040
Db	1981		
Qy	2041	CAGACCTCTGTCGGCTGCAAGCGCCCCCGGGCGCGTTCTTTGACACCATGACCTGC	2100
Db	2041		
Qy	2101	GGAGCTCGGACCTTTACTTGGTCAAGGATGCGGATGTCAITTCGGTGGCGCGGG	2160
Db	2101		
Qy	2161	GGGACAGCAGCGGGGAGCCTACTCTCCCGCAGGCGCGTCTCTACTTGAAGGCTCTTCG	2220
Db	2161		
Qy	2221	GGCGTCCACTGCTCTGCGCTCGGGCAGCGCTGTGGGCACTTTTCGGGCTGCGGTGCG	2280
Db	2221		
Qy	2281	ACCGAGGGTTGCGAGGGCGGTGGACTTTGTACCCGTCGAGTCTATGGAACCACTATG	2340
Db	2281		
Qy	2341	CGGTCCCGCTTCTCAGCGCAAACTCGTCCCTCGGCGCTACCGCAGACATTCAGGTG	2400
Db	2341		
Qy	2401	GCCCATCTACAGCGCTTCTGTTAGCGCAAGAGCACTAAGGTGCGGCTGCGTATGCA	2460
Db	2401		
Qy	2461	GCCCAAGGTTATAGGTGCTTCTGTAACCGCTCGTGGCGGCACTTAGGTTTCGG	2520
Db	2461		
Qy	2521	GCCTATATGTTAAGGCACATGTTATCGACCTCAACATCAGAACGGGTAAAGCAATC	2580
Db	2521		
Qy	2581	ACACGGGTGCGCCCATCAGTACTCCACTATGGCAAGTTCTTGGCGACGGTGTTCG	2640
Db	2581		
Qy	2641	TCTGGGCGCTTATGATCATATATATGATGAGTGCCTCACTCACTGACTGACCACT	2700
Db	2641		
Qy	2701	ATCTGGGCACTCGGCACTCTCTGGAACAGCGAGCGCTGAGCGGCTGCTGCTG	2760
Db	2701		
Qy	2761	CTGCCACCGCTACGCTCGGGATCGGTACCGTGCACATCCAAACATCGAGGAGGTG	2820
Db	2761		
Qy	2821	GCTCTGTCAGCTAGGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC	2880
Db	2821		
Qy	2881	AAGGGGGGAGGACCTCTATTTCTGCCATTCACAAAGAAATGTATGATGCTGCGCGG	2940
Db	2881		











Db 3181 CAGCGCGAGGCGAGCTGGTAGGGCAGAGTGGGCAATTTACAGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGGCCCTCGGGCATGTTGCAATTCCTCGGTTCTGTGGAGTGTATGACGGGCTGT 3300  
Db 3241 GAACGGCCCTCGGGCATGTTGCAATTCCTCGGTTCTGTGGAGTGTATGACGGGCTGT 3300  
QY 3301 GCTTGTAGAGCTCAGCGCCGCGAGACCTCAGTTAGTTGGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTAGAGCTCAGCGCCGCGAGACCTCAGTTAGTTGGGGCTTACCTAAACACA 3360  
QY 3361 CAGAGTTGCGGCTGCGCAGGACCATCTGGAGTTCTGGAGAGCGCTCTTTACAGGCTC 3420  
Db 3361 CAGAGTTGCGGCTGCGCAGGACCATCTGGAGTTCTGGAGAGCGCTCTTTACAGGCTC 3420  
QY 3421 ACCCATAGAGCCCAATTTCTTGTCCAGACTTAAGCAGCGAGAGACAATTTCCCTTAC 3480  
Db 3421 ACCCATAGAGCCCAATTTCTTGTCCAGACTTAAGCAGCGAGAGACAATTTCCCTTAC 3480  
QY 3481 CTGGTAGCATACAGGCTAGGTTGCGCAGGCTCAGGCTCAGCTCCACCTCATCGTGGGAC 3540  
Db 3481 CTGGTAGCATACAGGCTAGGTTGCGCAGGCTCAGGCTCAGCTCCACCTCATCGTGGGAC 3540  
QY 3541 CAAATGTGAAGTGTCTCATACGGCTAAAGCCCTACGCTGACGGGCCAAGCCCTGCTG 3600  
Db 3541 CAAATGTGAAGTGTCTCATACGGCTAAAGCCCTACGCTGACGGGCCAAGCCCTGCTG 3600  
QY 3601 TATAGCTGGAGCCCTTCAAACAGAGTTTACTACACACACCCCATTAACCAATACATC 3660  
Db 3601 TATAGCTGGAGCCCTTCAAACAGAGTTTACTACACACACCCCATTAACCAATACATC 3660  
QY 3661 ATGGCATGATGTGCGCTGACCTGAGGTCGTCAACGACACCTGGTGTGTTAGCGGA 3720  
Db 3661 ATGGCATGATGTGCGCTGACCTGAGGTCGTCAACGACACCTGGTGTGTTAGCGGA 3720  
QY 3721 GTCTTAGCAGCTCTGCGCGGTATGCTGACAAACAGGAGCGCTGTGTTGGGCGAGG 3780  
Db 3721 GTCTTAGCAGCTCTGCGCGGTATGCTGACAAACAGGAGCGCTGTGTTGGGCGAGG 3780  
QY 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTCGCGACAGGGAATGCTTTACCGGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCGGCGCATCATTCGCGACAGGGAATGCTTTACCGGGAGTTC 3840  
QY 3841 GATGAGATGGAAGTGGCTCAGACCTCCCTTACATCGAAGAGGAAATGACGCTGCC 3900  
Db 3841 GATGAGATGGAAGTGGCTCAGACCTCCCTTACATCGAAGAGGAAATGACGCTGCC 3900  
QY 3901 GAACAAATTCAGAAAGCGGCGCATCATTCGCGACAGGGAATGCTTTACCGGGAGTTC 3960  
Db 3901 GAACAAATTCAGAAAGCGGCGCATCATTCGCGACAGGGAATGCTTTACCGGGAGTTC 3960  
QY 3961 GCTGCTCCGTTGTGGAATCCAGTGGCGAACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCGTTGTGGAATCCAGTGGCGAACCTCGAAGCCTTCTGGGCGAAGCATATG 4020  
QY 4021 TGAATTTATCAGCGGATACAAATTTAGAGGCTTGTCCATCTGCTGCTGCAACCCC 4080  
Db 4021 TGAATTTATCAGCGGATACAAATTTAGAGGCTTGTCCATCTGCTGCTGCAACCCC 4080  
QY 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCCCGCTCACCACCAACAT 4140  
Db 4081 GCGATAGCATCACTGATGGCAATTCACAGCCTCTATCACAGCCCGCTCACCACCAACAT 4140  
QY 4141 ACCCTCTGTTTAACTCTGGGGGATGGTGGCGCCCAACTGCTGCTCTCCAGGCT 4200  
Db 4141 ACCCTCTGTTTAACTCTGGGGGATGGTGGCGCCCAACTGCTGCTCTCCAGGCT 4200  
QY 4201 GCTTCTGCTTTGTAGCGCGGCTCAGCTGGAGCGGCTGTGGAGCATAGGCTTTGGG 4260  
Db 4201 GCTTCTGCTTTGTAGCGCGGCTCAGCTGGAGCGGCTGTGGAGCATAGGCTTTGGG 4260  
QY 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGAGGGTGGCAGCGCGCTCGTGCC 4320

Db 4261 AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGCGCGCTCGTGCC 4320  
QY 4321 TTTTAAAGTCAACAGCGGCGAGATGCCCTCACCGAGGACCTGTTAACTACTCCCTGCT 4380  
Db 4321 TTTTAAAGTCAACAGCGGCGAGATGCCCTCACCGAGGACCTGTTAACTACTCCCTGCT 4380  
QY 4381 ATCTCTCTCCCTCGGGCCCTAGTCTCGGGTCTGTGTGCGCAGCGATCTGCTGCGGCAC 4440  
Db 4381 ATCTCTCTCCCTCGGGCCCTAGTCTCGGGTCTGTGTGCGCAGCGATCTGCTGCGGCAC 4440  
QY 4441 GTGGGCCACAGGGAGGGGCTGTGAGTGGATGAACCCGCTGATAGCGTTGCTTCCGG 4500  
Db 4441 GTGGGCCACAGGGAGGGGCTGTGAGTGGATGAACCCGCTGATAGCGTTGCTTCCGG 4500  
QY 4501 GGTAAACAGCTCTCCCCACGACATGTCCTGAGAGCGCTCAGCAGCATGTCACCT 4560  
Db 4501 GGTAAACAGCTCTCCCCACGACATGTCCTGAGAGCGCTCAGCAGCATGTCACCT 4560  
QY 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCACAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTCTAGTCTTACCATCACTCAGCTGCTGAAGAGGCTTCACAGTGGATCAAC 4620  
QY 4621 GAGGAGTGTCTCAGCCATGCTCCGGCTCGTGGCTGAAGAGTGTGGGATTTGGATATGC 4680  
Db 4621 GAGGAGTGTCTCAGCCATGCTCCGGCTCGTGGCTGAAGAGTGTGGGATTTGGATATGC 4680  
QY 4681 ACGGTTGCTGACATTTCAAGACCTGCTCAGTCCAGCTCCTGCGCGGATTTGCCGGA 4740  
Db 4681 ACGGTTGCTGACATTTCAAGACCTGCTCAGTCCAGCTCCTGCGCGGATTTGCCGGA 4740  
QY 4741 GTCCCTCTCTCTCAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 4800  
Db 4741 GTCCCTCTCTCTCAATGTCATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 4800  
QY 4801 CAAACACCTGCGCATGTCGAGCAGATCAGCGCATGTCGAGGAGTTCGAGGAGTTCATGAGG 4860  
Db 4801 CAAACACCTGCGCATGTCGAGCAGATCAGCGCATGTCGAGGAGTTCGAGGAGTTCATGAGG 4860  
QY 4861 ATCTGCGGCGCTAGGACCTGTAGTAAACGTCATGTCGAGGAGTTCGAGGAGTTCATGAGG 4920  
Db 4861 ATCTGCGGCGCTAGGACCTGTAGTAAACGTCATGTCGAGGAGTTCGAGGAGTTCATGAGG 4920  
QY 4921 ACCACGCGCGCTGACGCGCTCCTCCGCGCGCAATTTATTTAGGCGCTGTGCGGCGT 4980  
Db 4921 ACCACGCGCGCTGACGCGCTCCTCCGCGCGCAATTTATTTAGGCGCTGTGCGGCGT 4980  
QY 4981 GCTGCTGAGAGTACGTCGAGGTTACGCGGCTGCGGAGTTCACACTACGTCGCGGCGATG 5040  
Db 4981 GCTGCTGAGAGTACGTCGAGGTTACGCGGCTGCGGAGTTCACACTACGTCGCGGCGATG 5040  
QY 5041 ACCACTGACAACTAAAGTCCCGCTGTAGGTTCCGCGCGCAATTTCTTTCACAGAGT 5100  
Db 5041 ACCACTGACAACTAAAGTCCCGCTGTAGGTTCCGCGCGCAATTTCTTTCACAGAGT 5100  
QY 5101 GATGCGGCTGCGGTTGACAGGTTACGCTCAGCGTGCAGGAGTTCCTTACCGGAGGAGT 5160  
Db 5101 GATGCGGCTGCGGTTGACAGGTTACGCTCAGCGTGCAGGAGTTCCTTACCGGAGGAGT 5160  
QY 5161 ACATTCCTGCTGGGCTCAATCAATACCTGCTGAGGTTACAGCTCCCATCGAGCGCGAA 5220  
Db 5161 ACATTCCTGCTGGGCTCAATCAATACCTGCTGAGGTTACAGCTCCCATCGAGCGCGAA 5220  
QY 5221 CCGGAGTACAGTGTCTTCACTTCCAGTCCAGGAGTTCAGGCTCCAGGAGTTCAGGAGT 5280  
Db 5221 CCGGAGTACAGTGTCTTCACTTCCAGTCCAGGAGTTCAGGCTCCAGGAGTTCAGGAGT 5280  
QY 5281 GCTAAGCGTAGGCTGCGGAGGAGTTCCTGCTGCGGAGTTCAGGCTCCAGGAGTTCAGGAGT 5340  
Db 5281 GCTAAGCGTAGGCTGCGGAGGAGTTCCTGCTGCGGAGTTCAGGCTCCAGGAGTTCAGGAGT 5340  
QY 5341 CTGCTGCGGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGCTGAC 5400  
Db 5341 CTGCTGCGGCTTCTTGAAGGCAACATGCACTACCGCTCATGACTCCCGGAGCTGAC 5400

QY 5401 CTCATCGAGGCAACCTCTCTGTGGCGCAGAGATGGGCGGGAACATACCCGCGTGAG 5460  
DB 5401 CTCATCGAGGCAACCTCTCTGTGGCGCAGAGATGGGCGGGAACATACCCGCGTGAG 5460  
QY 5461 TCAGAAATAAGGTAGTAAATTTTGSACTCTTTTCGAGCCGCTCCAAAGCGGAGGAGATGAG 5520  
DB 5461 TCAGAAATAAGGTAGTAAATTTTGSACTCTTTTCGAGCCGCTCCAAAGCGGAGGAGATGAG 5520  
QY 5521 AGGGAAGTATCCGTTCGGCGGAGATCTCTGGGAGGTCCAGGAATTCCTTCGAGCGGATG 5580  
DB 5521 AGGGAAGTATCCGTTCGGCGGAGATCTCTGGGAGGTCCAGGAATTCCTTCGAGCGGATG 5580  
QY 5581 CCCATATGGGCACGCCCGGATTAACAACCTCTCACTGTTAGAGTCTCTGGAAGGACCCGAC 5640  
DB 5581 CCCATATGGGCACGCCCGGATTAACAACCTCTCACTGTTAGAGTCTCTGGAAGGACCCGAC 5640  
QY 5641 TAGCTCCTCCAGTGGTACAGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA 5700  
DB 5641 TAGCTCCTCCAGTGGTACAGGGTGTCCATTGCGCGCTGCCAAGGCCCTCCGATACCA 5700  
QY 5701 CCTTCCAGGAGGAGAGAGCGGTTGTCTGTCTAGAAATCTACCGTGTCTTCTGCTTGGCG 5760  
DB 5701 CCTTCCAGGAGGAGAGAGCGGTTGTCTGTCTAGAAATCTACCGTGTCTTCTGCTTGGCG 5760  
QY 5761 GAGCTCGCCACAAAGACCTTCGGCAGCTCCGAATCGTCGGCCGTCGACAGCGGCAACGCCA 5820  
DB 5761 GAGCTCGCCACAAAGACCTTCGGCAGCTCCGAATCGTCGGCCGTCGACAGCGGCAACGCCA 5820  
QY 5821 ACGGCTCTCTGTACACGAGCTCCGACGAGCGGCGGATCCGAGCTTGAGTCTGTAC 5880  
DB 5821 ACGGCTCTCTGTACACGAGCTCCGACGAGCGGCGGATCCGAGCTTGAGTCTGTAC 5880  
QY 5881 TCCTTCATGCCCTTTGAGGGGAGCGCGGGGATCCCGATCTCAGCGAGCGGTCTTTG 5940  
DB 5881 TCCTTCATGCCCTTTGAGGGGAGCGCGGGGATCCCGATCTCAGCGAGCGGTCTTTG 5940  
QY 5941 TCTACCGTAAAGGAGGAGTGTAGGAGCGTCTGTCTGCTGCTCGATCTCTACATG 6000  
DB 5941 TCTACCGTAAAGGAGGAGTGTAGGAGCGTCTGTCTGCTGCTCGATCTCTACATG 6000  
QY 6001 ACAGGCGCTGTATCAAGCCATGCGCTGCGGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
DB 6001 ACAGGCGCTGTATCAAGCCATGCGCTGCGGAGGAAACCAAGCTGCCATCAATGCACTG 6060  
QY 6061 AGCAACTCTTTGCTCGCTCACCACCTTGGTCTATGCTACAACTCTCGAGCGCAAGC 6120  
DB 6061 AGCAACTCTTTGCTCGCTCACCACCTTGGTCTATGCTACAACTCTCGAGCGCAAGC 6120  
QY 6121 CTGCGGAGAGAGGTACCTTTGACAGACTGAGGTCTGAGCGAGGAAACCAAGCTGCCATCAATGCACTG 6180  
DB 6121 CTGCGGAGAGAGGTACCTTTGACAGACTGAGGTCTGAGCGAGGAAACCAAGCTGCCATCAATGCACTG 6180  
QY 6181 GTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAAGCTAACTTCTATCCGTGGAG 6240  
DB 6181 GTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAAGCTAACTTCTATCCGTGGAG 6240  
QY 6241 GAAGCTGTAAAGTACGCGCCCAANTTCGGCGAGATCTAAATTTGGCTATGGGCAAG 6300  
DB 6241 GAAGCTGTAAAGTACGCGCCCAANTTCGGCGAGATCTAAATTTGGCTATGGGCAAG 6300  
QY 6301 GAGTCCGGAACCTATCCAGGAGGCGTTAAACCATTCGCTCCGCTGCGGAGGAGCTTG 6360  
DB 6301 GAGTCCGGAACCTATCCAGGAGGCGTTAAACCATTCGCTCCGCTGCGGAGGAGCTTG 6360  
QY 6361 CTGGAAGACACTGAGACACCAATTGACACCAACCATCATGCAAAATAAGAGTTTCTG 6420  
DB 6361 CTGGAAGACACTGAGACACCAATTGACACCAACCATCATGCAAAATAAGAGTTTCTG 6420  
QY 6421 GTCCAAACGAGAGGAGGCGGCAAGCCAGCTGCGCTTATCGTATTCGAGATTTGGG 6480  
DB 6421 GTCCAAACGAGAGGAGGCGGCAAGCCAGCTGCGCTTATCGTATTCGAGATTTGGG 6480

QY 6481 GTTCGTGTGTCGAGAGAAATGGCCCTTTTACGATGTGTTCTCAACCTCCCTCAGGCCGTG 6540  
DB 6481 GTTCGTGTGTCGAGAGAAATGGCCCTTTTACGATGTGTTCTCAACCTCCCTCAGGCCGTG 6540  
QY 6541 ATGGGCTCTTCATACGGAATTCGAATCTCTCTGACAGCGGTCGAGTCTCTGGTGAAT 6600  
DB 6541 ATGGGCTCTTCATACGGAATTCGAATCTCTCTGACAGCGGTCGAGTCTCTGGTGAAT 6600  
QY 6601 GCCTGGAAGCGAAGAAATGCCCTTATGGGCTTCGATATGACACCCGCTGTTTGTACTCA 6660  
DB 6601 GCCTGGAAGCGAAGAAATGCCCTTATGGGCTTCGATATGACACCCGCTGTTTGTACTCA 6660  
QY 6661 ACGGTCTACTGAGATGATCCGTTGTGAGGAGTCAATCTACCAATGTTGTGACTTGCC 6720  
DB 6661 ACGGTCTACTGAGATGATCCGTTGTGAGGAGTCAATCTACCAATGTTGTGACTTGCC 6720  
QY 6721 CCCGAAGCGAGAGGCGCATTAAGTCTCTACAGAGCGGCTTTACATCGGGGGCCCCCTG 6780  
DB 6721 CCCGAAGCGAGAGGCGCATTAAGTCTCTACAGAGCGGCTTTACATCGGGGGCCCCCTG 6780  
QY 6781 ACTAATTTTAAAGGCGAGAACTGCGGCTATGCGCGGTGCGCGGAGCGGTGTACTGACG 6840  
DB 6781 ACTAATTTTAAAGGCGAGAACTGCGGCTATGCGCGGTGCGCGGAGCGGTGTACTGACG 6840  
QY 6841 ACCAGCTGCGGTAATACCTTCAATGTTACTTGAAGCGCGCTGCGGCTGTGAGGCTGCG 6900  
DB 6841 ACCAGCTGCGGTAATACCTTCAATGTTACTTGAAGCGCGCTGCGGCTGTGAGGCTGCG 6900  
QY 6901 AAGCTTCAGGACTGTCAGATGCTCGTATGCGGAGACGACCTTGTGCTTATCTGTGAAGC 6960  
DB 6901 AAGCTTCAGGACTGTCAGATGCTCGTATGCGGAGACGACCTTGTGCTTATCTGTGAAGC 6960  
QY 6961 GCGGGACCCCAAGAGGAGCGGCGGCTTACGGGCTTTCAGCGAGGCTATGACTAGATAC 7020  
DB 6961 GCGGGACCCCAAGAGGAGCGGCGGCTTACGGGCTTTCAGCGAGGCTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTCTGGGAGCCGCCAAACCAAGATACGACTTGGAGTTGATAACATCATGC 7080  
DB 7021 TCTGCCCCCTCTGGGAGCCGCCAAACCAAGATACGACTTGGAGTTGATAACATCATGC 7080  
QY 7081 TCTTCCAACTGTTCAGTCGCGCAGATGCTATGCGGAGGAGGCTGACTATCTCACCCGT 7140  
DB 7081 TCTTCCAACTGTTCAGTCGCGCAGATGCTATGCGGAGGAGGCTGACTATCTCACCCGT 7140  
QY 7141 GACCCCAACCCCTTTCGCGGCTGCGTGGGAGACAGTACACACTCCAGTCAAT 7200  
DB 7141 GACCCCAACCCCTTTCGCGGCTGCGTGGGAGACAGTACACACTCCAGTCAAT 7200  
QY 7201 TCTGCGTAGGCAACATCATGATGCGCCCACTTGTGGGCAAGGATGATCTGATG 7260  
DB 7201 TCTGCGTAGGCAACATCATGATGCGCCCACTTGTGGGCAAGGATGATCTGATG 7260  
QY 7261 ACTCATTTCTTCCATCTTCTAGCTCAGGAAACAACTTGAAGGCTAGATGTCAG 7320  
DB 7261 ACTCATTTCTTCCATCTTCTAGCTCAGGAAACAACTTGAAGGCTAGATGTCAG 7320  
QY 7321 ATCTACGGGCTGTCTTCTCAATTTGAGCCACTTGAACCTCAGATCATCAACGACTC 7380  
DB 7321 ATCTACGGGCTGTCTTCTCAATTTGAGCCACTTGAACCTCAGATCATCAACGACTC 7380  
QY 7381 CATGGCTTAGGCGATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
DB 7381 CATGGCTTAGGCGATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGCT 7440  
QY 7441 TCATGCTCAGGAACTTGGGTTACCGCTTTCGAGTCTGGAGATCGGCGCAGAGT 7500  
DB 7441 TCATGCTCAGGAACTTGGGTTACCGCTTTCGAGTCTGGAGATCGGCGCAGAGT 7500  
QY 7501 GTCCGCGCTAGGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
DB 7501 GTCCGCGCTAGGCTACTGTCCAGGGGGGAGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
QY 7561 AACTGGGAGTAGGACCAAGCTCAAACTCACTCCAAATCCCGCTGCTGCCAGTTGGAT 7620

Db	7561	AACTGGGCAGTATAGGACCAAGCTCAAAACTCACTCCAAATCCCGGTGCTCCCAAGTTGGAT	7620
Qy	7621	TTATCCAGCTGGTTCTGTTGCTGGTTACAGCGGGGAGACATATACACGCTGCTCTCGT	7680
Db	7621	TTATCCAGCTGGTTCTGTTGCTGGTTACAGCGGGGAGACATATACACGCTGCTCTCGT	7680
Qy	7681	GCCGAGCCCGCTGGTTCTCATGTGGTGCTACTCCTACTCTTCTGTAGGGGTAGGCATCTAT	7740
Db	7681	GCCGAGCCCGCTGGTTCTCATGTGGTGCTACTCCTACTCTTCTGTAGGGGTAGGCATCTAT	7740
Qy	7741	CTACTCCCCAACCGATGAACGGGGACCTAAACACTCCAGGCCAATAGGCCATCCTGTTTT	7800
Db	7741	CTACTCCCCAACCGATGAACGGGGAGCTAAACACTCCAGGCCAATAGGCCATCCTGTTTT	7800
Qy	7801	TTTCCCTTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCCTTT	7860
Db	7801	TTTCCCTTTTTTTTTTTCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCTCCCTTT	7860
Qy	7861	TTTTTCTCTTTTTTTTCTCTTTCTTTCTTTGGTGGCTCCATCTTAGCCCTAGTCACGGC	7920
Db	7861	TTTTTCTCTTTTTTTTCTCTTTCTTTCTTTGGTGGCTCCATCTTAGCCCTAGTCACGGC	7920
Qy	7921	TAGCTGTCAAAAGTCCGTGAGCCGCTTGACTTCGACAGAGTGCTGATACTGSCCTCTCTGC	7980
Db	7921	TAGCTGTCAAAAGTCCGTGAGCCGCTTGACTTCGACAGAGTGCTGATACTGSCCTCTCTGC	7980
Qy	7981	AGATCAAGT	7989
Db	7981	AGATCAAGT	7989

RESULT 13  
 US-10-434-842-17  
 ; Sequence 17, Application US/10434842  
 ; Publication No. US20040005549A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPLICATION  
 ; FILE REFERENCE: 0342/1H395US3  
 ; CURRENT APPLICATION NUMBER: US/10/434,842  
 ; CURRENT FILING DATE: 2003-05-09  
 ; PRIOR APPLICATION NUMBER: US 10/233,307  
 ; PRIOR FILING DATE: 2002-08-28  
 ; PRIOR APPLICATION NUMBER: US 10/005,469  
 ; PRIOR FILING DATE: 2001-11-07  
 ; PRIOR APPLICATION NUMBER: US 60/245,866  
 ; PRIOR FILING DATE: 2000-11-07  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 17  
 ; LENGTH: 7992  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: FCA22 Replicon Sequence  
 US-10-434-842-17

	Query Match	Best Local Similarity	99.9%;	Score	7982.6;	DB	16;	Length	7992;	
	Matches	7985;	Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;
QY	1	GCACGCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTCTGAGGA	ACTCTG	60						
DB	1	GCACGCCCGATTGGGGGCGACACTCCACCATAGATCACTCCCTCTGAGGA	ACTCTG	60						
QY	61	TCCTTCACGCAGAAACGGCTCTAGCCATGGCGTTAGTATGAGTGTCTGTG	CAGCCTCCAGGAC	120						
DB	61	TCCTTCACGCAGAAACGGCTCTAGCCATGGCGTTAGTATGAGTGTCTGTG	CAGCCTCCAGGAC	120						
QY	121	CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAA	CCGGT	GAGTACACCGGAATTC	CCAG	180				
DB	121	CCCCCTCCCGGAGAGCCATAGTGGTCTGCGGAA	CCGGT	GAGTACACCGGAATTC	CCAG	180				

181	GACGACGGGTCTCTTCTTTGGATCAACCGCTCAATGCTTCGGAGATTGGCGCTGCCCC	240
181	GACGACGGGTCTCTTCTTTGGATCAACCGCTCAATGCTTCGGAGATTGGCGCTGCCCC	240
241	CGGAGACTGCTAGCCGAGTAGTGTGGGTGCGAAAGCCCTTGTGGTACTGCCCTGATAGG	300
241	CGGAGACTGCTAGCCGAGTAGTGTGGGTGCGAAAGCCCTTGTGGTACTGCCCTGATAGG	300
301	GTGCTTCGAGTCCCCCGGAGGTCTGTPAGACCGTGCAACATGAGCAGCAAGATCTCTAAAC	360
301	GTGCTTCGAGTCCCCCGGAGGTCTGTPAGACCGTGCAACATGAGCAGCAAGATCTCTAAAC	360
361	CTCAAGAAATAAACAAGGCGCGCCATGATTGAACAAGATGGAATGCACGCAAGTTCTC	420
361	CTCAAGAAATAAACAAGGCGCGCCATGATTGAACAAGATGGAATGCACGCAAGTTCTC	420
421	CGGCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTTGGGCAACAACAGCAATCGGTGCT	480
421	CGGCGCTTGGGTGGAGAGGCTATTTCGGCTATGACTTGGGCAACAACAGCAATCGGTGCT	480
481	CTGATGCGCGGTGTTCCGGCTGTGACGACGAGGCGCGCCGGTCTTTTGTCAAGACCG	540
481	CTGATGCGCGGTGTTCCGGCTGTGACGACGAGGCGCGCCGGTCTTTTGTCAAGACCG	540
541	ACCTGTCGGTGCCTGAATGAACCTCCAGACAAGGCAAGCGCGCTATCTGGTGTGGCCA	600
541	ACCTGTCGGTGCCTGAATGAACCTCCAGACAAGGCAAGCGCGCTATCTGGTGTGGCCA	600
601	CGACGGCGTTCCTTGGCGAGCTGTCTCGAGCTTGTCACTGAACGGGGAAGGACTGGC	660
601	CGACGGCGTTCCTTGGCGAGCTGTCTCGAGCTTGTCACTGAACGGGGAAGGACTGGC	660
661	TGCTATTGGGCGAAGTCCCGGCGCAGATCTCTCTGATCTCACTTGTCTCCTGCCGAGA	720
661	TGCTATTGGGCGAAGTCCCGGCGCAGATCTCTCTGATCTCACTTGTCTCCTGCCGAGA	720
721	AAGTATCCATCATGGCTGATGCAATCGCGCTGCATACGCTTGATCCGGCTACCTGCC	780
721	AAGTATCCATCATGGCTGATGCAATCGCGCTGCATACGCTTGATCCGGCTACCTGCC	780
781	CATTGACCCCAAGCAACATGCATCGACGGACGCTACTCGGATGGAACCGGCTC	840
781	CATTGACCCCAAGCAACATGCATCGACGGACGCTACTCGGATGGAACCGGCTC	840
841	TTGTGATCAGGATGATCTGGAAGAGAGCATCAGGGGTCCGCGCAGCCGAACTGTTCG	900
841	TTGTGATCAGGATGATCTGGAAGAGAGCATCAGGGGTCCGCGCAGCCGAACTGTTCG	900
901	CAAGCTCAAGCGCGGATGCGGACGGGAGATCTGCTGTGACCCATGGCGATGCTC	960
901	CAAGCTCAAGCGCGGATGCGGACGGGAGATCTGCTGTGACCCATGGCGATGCTC	960
961	GCTTCCGAATATCATGTTGMAAATGGCGCTTTTTCGGATTCATCGACTGTGCGCGGC	1020
961	GCTTCCGAATATCATGTTGMAAATGGCGCTTTTTCGGATTCATCGACTGTGCGCGGC	1020
1021	TGGGTGTGGCGAACCGCTATCAGACATAGCGTTGGCTACCGTGATATGCTGAGAGC	1080
1021	TGGGTGTGGCGAACCGCTATCAGACATAGCGTTGGCTACCGTGATATGCTGAGAGC	1080
1081	TTGGCGGGAATGGCTGACCGCTTCTGTTTACGGTATCGCGCTCCCGATTCGC	1140
1081	TTGGCGGCGAATGGGCTGACCGCTTCTGTTTACGGTATCGCGCTCCCGATTCGC	1140
1141	AGCGCATCGCTTCTATCGCCTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACAG	1200
1141	AGCGCATCGCTTCTATCGCCTCTTGACGAGTCTTCTGAGTTTAAACAGACCAACAG	1200
1201	GTTTCCCTCTAGCGGGATCAATCCGCCCTCTCCCTCCCGCGCCCTTAAAGTTACTGCG	1260
1201	GTTTCCCTCTAGCGGGATCAATCCGCCCTCTCCCTCCCGCGCCCTTAAAGTTACTGCG	1260



QY 1261 CGAACCGCTTGGAAATAAGCGCGGTGTGGTTGTCTATATGTTATTTTCCACCATATTG 1320  
DB 1261 CGAACCGCTTGGAAATAAGCGCGGTGTGGTTGTCTATATGTTATTTTCCACCATATTG 1320  
QY 1321 CCGTCTTTTGGCAATGTAGGGCCCGGAAAACCTGGCCCTGTCTTTTGAAGAGCATTCCT 1380  
DB 1321 CCGTCTTTTGGCAATGTAGGGCCCGGAAAACCTGGCCCTGTCTTTTGAAGAGCATTCCT 1380  
QY 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCTGTAAGAGCA 1440  
DB 1381 AGGGGTCTTTCCCTCTCGCCAAAGGAATGCAAGGTCTGTTGAATGTCTGTAAGAGCA 1440  
QY 1441 GTTCTCTGGAAGCTTTCTTGAAGACAAACAAAGCTCTGTAGGACCCCTTTGACGGCAGCGG 1500  
DB 1441 GTTCTCTGGAAGCTTTCTTGAAGACAAACAAAGCTCTGTAGGACCCCTTTGACGGCAGCGG 1500  
QY 1501 AACCCCCACCTGGCGCAGGTGCTCTGGGCGCAAAAGCCCAAGCTTTGACGGCAGCGG 1560  
DB 1501 AACCCCCACCTGGCGCAGGTGCTCTGGGCGCAAAAGCCCAAGCTTTGACGGCAGCGG 1560  
QY 1561 GCMAAGCGGCACAAACCCAGTGCACAGTGTGTAGTGGATAGTGTGGAAGAGTCAAA 1620  
DB 1561 GCMAAGCGGCACAAACCCAGTGCACAGTGTGTAGTGGATAGTGTGGAAGAGTCAAA 1620  
QY 1621 TGGTCTCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAAGGTACCCCATTTG 1680  
DB 1621 TGGTCTCTCAAGCGTATTCAACAAAGGGGCTGAAGGATGCCAAGGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTTACATGTGTTAGTCAAGGTTAAAA 1740  
DB 1681 ATGGGATCTGATCTGGGCGCTCGGTGCAATGCTTTTACATGTGTTAGTCAAGGTTAAAA 1740  
QY 1741 AACGTCTAGGCGCCCGGACCAACCGGAGCGTGTGTTTCCCTTTGAAAACACGATAATACC 1800  
DB 1741 AACGTCTAGGCGCCCGGACCAACCGGAGCGTGTGTTTCCCTTTGAAAACACGATAATACC 1800  
QY 1801 ATGGCGCTATTAGCGGCTACTCCCAACAGACGCGAGGCGCTTGGTGTGATCATCACT 1860  
DB 1801 ATGGCGCTATTAGCGGCTACTCCCAACAGACGCGAGGCGCTTGGTGTGATCATCACT 1860  
QY 1861 AGCTCTCAGCGCGGACAGGAAACAGGTGAGGGGAGGTCCTCAAGTGGTCTCCACCGCA 1920  
DB 1861 AGCTCTCAGCGCGGACAGGAAACAGGTGAGGGGAGGTCCTCAAGTGGTCTCCACCGCA 1920  
QY 1921 ACAAAATCTTCTGCGGACCTGCTCAATGGCGGTGTGTTGGAGTGTCTATCATGTGGCC 1980  
DB 1921 ACAAAATCTTCTGCGGACCTGCTCAATGGCGGTGTGTTGGAGTGTCTATCATGTGGCC 1980  
QY 1981 GGCTCAAAGACCTTTGCGGCGCCAAAGGGCCCAATCAACCAAAATGTACACCAATGTGGAC 2040  
DB 1981 GGCTCAAAGACCTTTGCGGCGCCAAAGGGCCCAATCAACCAAAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGTTCTTTGACACCAATGCACCTGC 2100  
DB 2041 CAGGACCTCGTGGCTGGCAAGCGCCCGCGGGCGGTTCTTTGACACCAATGCACCTGC 2100  
QY 2101 GGCAGCTCGGACCTTTACTTGTGTCACGAGCATGCGGATGCTATTCGGGTGCGCGCGG 2160  
DB 2101 GGCAGCTCGGACCTTTACTTGTGTCACGAGCATGCGGATGCTATTCGGGTGCGCGCGG 2160  
QY 2161 GGCAGACAGGGGGAGGCTACTTCTCCCGCAGGCGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
DB 2161 GGCAGACAGGGGGAGGCTACTTCTCCCGCAGGCGCGCTCTCTACTTTGAAGGGCTCTTCG 2220  
QY 2221 GCGGTCTCACTGCTCTGCGGCGACGCTGTGGGCAATCTTTTCGGGCTGCGGTGTGC 2280  
DB 2221 GCGGTCTCACTGCTCTGCGGCGACGCTGTGGGCAATCTTTTCGGGCTGCGGTGTGC 2280  
QY 2281 ACCCGAGGGGTTCGAAGGGCGGTGGAATTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
DB 2281 ACCCGAGGGGTTCGAAGGGCGGTGGAATTTGTACCCGTGAGTCTATGGAACCACTATG 2340  
QY 2341 CCGTCCCGGCTTTCACGGCAACTCGTCCCTCCGCGCGTACCGCAGACATTCACGGTG 2400

DB 2341 CCGTCCCGGCTTTCACGGCAACTCGTCCCTCCGCGCGTACCCGACACATTCACGGTG 2400  
QY 2401 GCCATCTACACGCGCCCTACTGGTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
DB 2401 GCCATCTACACGCGCCCTACTGGTAGCGGCAAGACACTAAGGTGCGGCTCGGTATGCA 2460  
QY 2461 GCCAAGGGTATAGGGTCTTGTCTGAAACCGCTCCGTCCGCGCCACCTAGTTTCGGG 2520  
DB 2461 GCCAAGGGTATAGGGTCTTGTCTGAAACCGCTCCGTCCGCGCCACCTAGTTTCGGG 2520  
QY 2521 CGGTATATGTCTAAGGCACATGTTATCGACCTTACATCAGAACCGGGTAAAGACCATC 2580  
DB 2521 CGGTATATGTCTAAGGCACATGTTATCGACCTTACATCAGAACCGGGTAAAGACCATC 2580  
QY 2581 ACCACGGGTGCCCCCATCACTACTCCACCTATGCAAGTTCCTGCGCAGCGGTGTTGC 2640  
DB 2581 ACCACGGGTGCCCCCATCACTACTCCACCTATGCAAGTTCCTGCGCAGCGGTGTTGC 2640  
QY 2641 TCTGGGGCGGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
DB 2641 TCTGGGGCGGCTATGACATCATATATGTGATGAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTGTGGCATCGGCACAGTCTCGSACCAAGCGGAGCGGCTGAGCGGACTCGTCTG 2760  
DB 2701 ATCTGTGGCATCGGCACAGTCTCGSACCAAGCGGAGCGGCTGAGCGGACTCGTCTG 2760  
QY 2761 CTCGCCACCGCTACGCTTCCGGGATCGGTCAACCTGCGCACATCCAAACATCGAGAGTG 2820  
DB 2761 CTCGCCACCGCTACGCTTCCGGGATCGGTCAACCTGCGCACATCCAAACATCGAGAGTG 2820  
QY 2821 GCTCTGTCCAGCATCGAGAAATCCCTTTTATGGCAAGCCATCCCATCCAGACCATC 2880  
DB 2821 GCTCTGTCCAGCATCGAGAAATCCCTTTTATGGCAAGCCATCCCATCCAGACCATC 2880  
QY 2881 AAGGGGGGAGGCGACCTCATTTCTGCGCATTCGAAGAAATGTGATGAGCTCGCGCG 2940  
DB 2881 AAGGGGGGAGGCGACCTCATTTCTGCGCATTCGAAGAAATGTGATGAGCTCGCGCG 2940  
QY 2941 AAGGTCTCCGGCTTCGACCTCAATGCTGTAGCATATTACCGGGGCTTGTATGTTCCGTC 3000  
DB 2941 AAGGTCTCCGGCTTCGACCTCAATGCTGTAGCATATTACCGGGGCTTGTATGTTCCGTC 3000  
QY 3001 ATACCACTAGCGGAGACGTCATTTGCTAGCAACGAGCGCTTAATGACGGGCTTTACC 3060  
DB 3001 ATACCACTAGCGGAGACGTCATTTGCTAGCAACGAGCGCTTAATGACGGGCTTTACC 3060  
QY 3061 GGGATTTGACACTCAGTGTGCAATGCTGTCATGTCACCCAGACAGTGCAGCTTCAGC 3120  
DB 3061 GGGATTTGACACTCAGTGTGCAATGCTGTCATGTCACCCAGACAGTGCAGCTTCAGC 3120  
QY 3121 CTGGACCCGACCTTACCATTTGAGACGACCGCTGCCAAGACGCGGTGTACGCTCG 3180  
DB 3121 CTGGACCCGACCTTACCATTTGAGACGACCGCTGCCAAGACGCGGTGTACGCTCG 3180  
QY 3181 CAGCGCGAGGCGAGGACTGTTAGGGCAGGATGGGCAATTTACAGTTTGTGACTCCAGGA 3240  
DB 3181 CAGCGCGAGGCGAGGACTGTTAGGGCAGGATGGGCAATTTACAGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGCGCTTCGGGATGTTTCGATTCCTCGTCTGTCGAGTGTCTATGACGCGGGTGT 3300  
DB 3241 GAACGCGCTTCGGGATGTTTCGATTCCTCGTCTGTCGAGTGTCTATGACGCGGGTGT 3300  
QY 3301 GCTTGTGTACGAGCTCACGCGCCCGGAGACTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
DB 3301 GCTTGTGTACGAGCTCACGCGCCCGGAGACTCAGTTAGGTTGCGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTGCCCTCTGCCAGGACCATCTGGAGTCTTGGAGAGCGTCTTTTACAGGCTC 3420  
DB 3361 CCAGGTTGCCCTCTGCCAGGACCATCTGGAGTCTTGGAGAGCGTCTTTTACAGGCTC 3420  
QY 3421 ACCCATAGACGCCCATTTCTTGTCCAGACTTAAGACGCGAGGACAACTCCCTAC 3480

Db	3421	ACCCACATAGAGCCCAATTTCTTGTCCTCCAGACTAAGAGCGGAGAGACAACATTCCTCCCTAC	3480	Qy	4561	CAGATCCTCTCTAGTCTTACCATCACTCAGCTGTCTGAAGAGGCTTCAACAGTGGATCAAC	4620
Qy	3481	CTGGTAGCATACAGGCTACGCTGTGCGCCAGGGCTCAGGCTCCACTCCATCTGTGGAC	3540	Db	4561	CAGATCCTCTCTAGTCTTACCATCACTCAGCTGTCTGAAGAGGCTTCAACAGTGGATCAAC	4620
Db	3481	CTGGTAGCATACAGGCTACGCTGTGCGCCAGGGCTCAGGCTCCACTCCATCTGTGGAC	3540	Qy	4621	GAGGACTGCTCCACGCCATGTCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATATGC	4680
Qy	3541	CAAATGTGAAGTGTCTCATACGGCTAAAGCCCTAAGCTGCACGGGCAACGCCCTGTGCTG	3600	Db	4621	GAGGACTGCTCCACGCCATGTCTCCGGCTCGTGGCTAAGAGATGTTTGGATTGGATATGC	4680
Db	3541	CAAATGTGAAGTGTCTCATACGGCTAAAGCCCTAAGCTGCACGGGCAACGCCCTGTGCTG	3600	Qy	4681	ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGTCTTCTGCGCGGATTTGCCGGGA	4740
Qy	3601	TATAGGCTGGAGCCGTTCAAAACAGAGTTTACTACACACCCCATACCAATATACATC	3660	Db	4681	ACGGTGTGACTGATTTCAAGACCTGGCTCCAGTCCAAAGTCTTCTGCGCGGATTTGCCGGGA	4740
Db	3601	TATAGGCTGGAGCCGTTCAAAACAGAGTTTACTACACACCCCATACCAATATACATC	3660	Qy	4741	GTCCCTCTTCTCATGTCAACGTGGTACAAAGGAGTCTGGCGGGGCGACGSCATCATG	4800
Qy	3661	ATGGCATGATGTGCGCTGACCTGGAGTGTGTCAAGACACCTGGTGTGTGTAGGCGGA	3720	Db	4741	GTCCCTCTTCTCATGTCAACGTGGTACAAAGGAGTCTGGCGGGGCGACGSCATCATG	4800
Db	3661	ATGGCATGATGTGCGCTGACCTGGAGTGTGTCAAGACACCTGGTGTGTGTAGGCGGA	3720	Qy	4801	CAAAACACCTGCCCCATGTGGAGCAGATCACCGGACATGTGAAACAAACGTTTCCATGAGG	4860
Qy	3721	GTCTAGCAGCTCTGGCCGCTATTGCTGTGACAAACAGGAGCGTGTCTATTGTGGGACG	3780	Db	4801	CAAAACACCTGCCCCATGTGGAGCAGATCACCGGACATGTGAAACAAACGTTTCCATGAGG	4860
Db	3721	GTCTAGCAGCTCTGGCCGCTATTGCTGTGACAAACAGGAGCGTGTCTATTGTGGGACG	3780	Qy	4861	ATCGTGGGCTTAGGACCTGTGTAGTAACACGTGGCATGGAAACATTCCTCCATTTAACGGGTAC	4920
Qy	3781	ATCATCTTGTCCGGAAGCCGSCATCATTCCTCCGACAGGGAAGTCTTTACCGGGAGTTC	3840	Db	4861	ATCGTGGGCTTAGGACCTGTGTAGTAACACGTGGCATGGAAACATTCCTCCATTTAACGGGTAC	4920
Db	3781	ATCATCTTGTCCGGAAGCCGSCATCATTCCTCCGACAGGGAAGTCTTTACCGGGAGTTC	3840	Qy	4921	ACCACGGGCTTGCACGCGCTCCCGCGGCCAAATTTATTCTAGGGCGCTGTGGCGGGTG	4980
Qy	3841	GATGAGATGGAAGTGCCTCTACACCTCCCTTATCATCGAACAGGGAATGCAGCTCGCC	3900	Db	4921	ACCACGGGCTTGCACGCGCTCCCGCGGCCAAATTTATTCTAGGGCGCTGTGGCGGGTG	4980
Db	3841	GATGAGATGGAAGTGCCTCTACACCTCCCTTATCATCGAACAGGGAATGCAGCTCGCC	3900	Qy	4981	GCTGCTGAGGAGTACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGCGGCATG	5040
Qy	3901	GAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGCT	3960	Db	4981	GCTGCTGAGGAGTACGTGAGGTTACGCGGTGGGGATTTCCACTACGTGACGCGGCATG	5040
Db	3901	GAACAATTCACACAGAGGCAATCGGTTGCTGCAACAGCCACCAAGCAAGCGGAGGCT	3960	Qy	5041	ACCACGTACAACTAAAGTCCCGTGTACAGTTCCGGCCCCCGAAATTCCTTCACAGAGTG	5100
Qy	3961	GCTGCTCCGTTGGTGAATCCAGTGGGAGCCCTCGAAGCTTCTGGGCGAAGCATATG	4020	Db	5041	ACCACGTACAACTAAAGTCCCGTGTACAGTTCCGGCCCCCGAAATTCCTTCACAGAGTG	5100
Db	3961	GCTGCTCCGTTGGTGAATCCAGTGGGAGCCCTCGAAGCTTCTGGGCGAAGCATATG	4020	Qy	5101	GATGGGTGCGGTGCACAGGTACGCTCCAGCGTGCACAAACCCCTCTACGGGAGGAGTTC	5160
Qy	4021	TGGAATTCATACGCGGATACAAATTTAGCAGGCTTGTCCATCTGCTGCGCAACCCC	4080	Db	5101	GATGGGTGCGGTGCACAGGTACGCTCCAGCGTGCACAAACCCCTCTACGGGAGGAGTTC	5160
Db	4021	TGGAATTCATACGCGGATACAAATTTAGCAGGCTTGTCCATCTGCTGCGCAACCCC	4080	Qy	5161	ACATTCCTGCTGGGTCAATCAATACCTTGGTTCACAGCTCCCATCGGAGCCGAA	5220
Qy	4081	GCATAGCATCACTGATGCATTCACAGCTCTATCACAGCCGCTCACCACCCACAT	4140	Db	5161	ACATTCCTGCTGGGTCAATCAATACCTTGGTTCACAGCTCCCATCGGAGCCGAA	5220
Db	4081	GCATAGCATCACTGATGCATTCACAGCTCTATCACAGCCGCTCACCACCCACAT	4140	Qy	5221	CCGAGCTAGCAGTGTCTCATCTCCATGCTCACGACCCCTCCACATTCACGGCGGAGACG	5280
Qy	4141	ACCTCTCTGTTAAACATCTCTGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGCGCT	4200	Db	5221	CCGAGCTAGCAGTGTCTCATCTCCATGCTCACGACCCCTCCACATTCACGGCGGAGACG	5280
Db	4141	ACCTCTCTGTTAAACATCTCTGGGGGATGGTGGCGCCCAACTTGTCTCTCCAGCGCT	4200	Qy	5281	GCTAAGCTAGGCTGGCGAGGATCTCCCTCTCTTGGCCAGCTCATCAGCTAGCCAG	5340
Qy	4201	GCTTCTGCTTGTAGCGCGCGGATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG	4260	Db	5281	GCTAAGCTAGGCTGGCGAGGATCTCCCTCTCTTGGCCAGCTCATCAGCTAGCCAG	5340
Db	4201	GCTTCTGCTTGTAGCGCGCGGATCGCTGGAGCGGCTGTTGGCAGCATAGGCTTGGG	4260	Qy	5341	CTGCTGTGCGCTTCTTGAAGGCAACATGACCTACCCGCTCATGACTCCCGGAGCGTGAC	5400
Qy	4261	AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGCGGCTCGTGCC	4320	Db	5341	CTGCTGTGCGCTTCTTGAAGGCAACATGACCTACCCGCTCATGACTCCCGGAGCGTGAC	5400
Db	4261	AAGGTGCTTGTGGATATTTTGGCAGGTTATGGAGCAGGGTGGCAGGCGGCTCGTGCC	4320	Qy	5401	CTCATCGAGCCAACTCTCTGTGGCGGAGATGGCGGGGAACATCACCCGCTGGAG	5460
Qy	4321	TTTAAAGTCAATAGCGCGGAGATGCCCTCCACCGAGGACCTGGTTAAACCTACTCCCTGCT	4380	Db	5401	CTCATCGAGCCAACTCTCTGTGGCGGAGATGGCGGGGAACATCACCCGCTGGAG	5460
Db	4321	TTTAAAGTCAATAGCGCGGAGATGCCCTCCACCGAGGACCTGGTTAAACCTACTCCCTGCT	4380	Qy	5461	TCAGAAATTAAGGTAGTAATTTTGGACTCTTTTCGAGCCGCTCCAAAGCGGAGGATGAG	5520
Qy	4381	ATCCTCTCCCTTGGCGGCTGTGCTGGGGTGTGTGCGAGCGATACCTGCTCGGCAC	4440	Db	5461	TCAGAAATTAAGGTAGTAATTTTGGACTCTTTTCGAGCCGCTCCAAAGCGGAGGATGAG	5520
Db	4381	ATCCTCTCCCTTGGCGGCTGTGCTGGGGTGTGTGCGAGCGATACCTGCTCGGCAC	4440	Qy	5521	AGGAAAGTATCCGTTCCGGCGGAGTCTCTCGGAGGTCCAGGAAATTCCTCGAGCGATG	5580
Qy	4441	GTGGGCCCCAGGGAGGGGCTGTGCACTGTGATGAACCGGCTGATAGCGTTTCGTTTCGGG	4500	Db	5521	AGGAAAGTATCCGTTCCGGCGGAGTCTCTCGGAGGTCCAGGAAATTCCTCGAGCGATG	5580
Db	4441	GTGGGCCCCAGGGAGGGGCTGTGCACTGTGATGAACCGGCTGATAGCGTTTCGTTTCGGG	4500	Qy	5581	CCCATATGGGACGCGCCCGGATTAACACCTTCTGTTAGTCTCTGGAAGACCCCGAC	5640
Qy	4501	GGTAAACACGCTCTCCCGACGCACTATGTGCTCGAGAGCGAGCTGAGCACTGTCACT	4560	Db	5581	CCCATATGGGACGCGCCCGGATTAACACCTTCTGTTAGTCTCTGGAAGACCCCGAC	5640
Db	4501	GGTAAACACGCTCTCCCGACGCACTATGTGCTCGAGAGCGAGCTGAGCACTGTCACT	4560				





QY 1561 GCAAGGCGGCACAAACCCAGTGCACAGTCTGTGAGTTGATGTAAGTCAAA 1620  
DB 1561 GCAAGGCGGCACAAACCCAGTGCACAGTCTGTGAGTTGATGTAAGTCAAA 1620  
QY 1621 TGGCTCTCCCTCAAGGCTATTCAACAAAGGGGCTGAAGGATGCCAAGAGTACCCCATTTG 1680  
DB 1621 TGGCTCTCCCTCAAGGCTATTCAACAAAGGGGCTGAAGGATGCCAAGAGTACCCCATTTG 1680  
QY 1681 ATGGGATCTGATCTGGGSCCTCGGTGCACATGCTTTTACATGTTTGTAGTCAGGTTAAAA 1740  
DB 1681 ATGGGATCTGATCTGGGSCCTCGGTGCACATGCTTTTACATGTTTGTAGTCAGGTTAAAA 1740  
QY 1741 AACGTCTAGGCCCCCGAAACACAGGAGAGTGGTTTTCCTTTGAAACACAGATAATACC 1800  
DB 1741 AACGTCTAGGCCCCCGAAACACAGGAGAGTGGTTTTCCTTTGAAACACAGATAATACC 1800  
QY 1801 ATGGCGCTATTACGGGCTATCTCCAAACAGACGGAGGCCCTACTTGGCTGCATCACT 1860  
DB 1801 ATGGCGCTATTACGGGCTATCTCCAAACAGACGGAGGCCCTACTTGGCTGCATCACT 1860  
QY 1861 AGCCTCACAGCGGGGACAGAACAGGTCCAGGGGAGGTCCAAAGTGGTCTCCACGCA 1920  
DB 1861 AGCCTCACAGCGGGGACAGAACAGGTCCAGGGGAGGTCCAAAGTGGTCTCCACGCA 1920  
QY 1921 ACACAAATCTTTCTGGGACCTCGGTCAAATGGCGTGTGGAGTGTCTATCATGTGGC 1980  
DB 1921 ACACAAATCTTTCTGGGACCTCGGTCAAATGGCGTGTGGAGTGTCTATCATGTGGC 1980  
QY 1981 GGCTCAAAGACCTTTGGCGGCCAAAGGGCCCAATCAACCAATGTACACCAATGTGGAC 2040  
DB 1981 GGCTCAAAGACCTTTGGCGGCCAAAGGGCCCAATCAACCAATGTACACCAATGTGGAC 2040  
QY 2041 CAGGACCTGTGGCTGCGGACCTTTACTTGGTCAACAGGATGCCAGTGTCACTTGAAGGGCTCTTCG 2100  
DB 2041 CAGGACCTGTGGCTGCGGACCTTTACTTGGTCAACAGGATGCCAGTGTCACTTGAAGGGCTCTTCG 2100  
QY 2101 GGCAGCTCGGACCTTTACTTGGTCAACAGGATGCCAGTGTCACTTGAAGGGCTCTTCG 2160  
DB 2101 GGCAGCTCGGACCTTTACTTGGTCAACAGGATGCCAGTGTCACTTGAAGGGCTCTTCG 2160  
QY 2161 GGCAGCTCGGACCTTTACTTGGTCAACAGGATGCCAGTGTCACTTGAAGGGCTCTTCG 2220  
DB 2161 GGCAGCTCGGACCTTTACTTGGTCAACAGGATGCCAGTGTCACTTGAAGGGCTCTTCG 2220  
QY 2221 GGCAGCTCGGACCTTTACTTGGTCAACAGGATGCCAGTGTCACTTGAAGGGCTCTTCG 2280  
DB 2221 GGCAGCTCGGACCTTTACTTGGTCAACAGGATGCCAGTGTCACTTGAAGGGCTCTTCG 2280  
QY 2281 ACCGAGGGGTTCGAGGCGGTGGAATTTGTACCCGTGAGTCTATGGAACCCACATATG 2340  
DB 2281 ACCGAGGGGTTCGAGGCGGTGGAATTTGTACCCGTGAGTCTATGGAACCCACATATG 2340  
QY 2341 CGGTCCCGGTCTTACGGAACAATCGTCCCTCCGCGGTACCGCAGACATTCAGAGTG 2400  
DB 2341 CGGTCCCGGTCTTACGGAACAATCGTCCCTCCGCGGTACCGCAGACATTCAGAGTG 2400  
QY 2401 GCCCATCTACACGCCCTTACTGTAGCGGCAAGAGACATTAAGTTCGGGTGATGCA 2460  
DB 2401 GCCCATCTACACGCCCTTACTGTAGCGGCAAGAGACATTAAGTTCGGGTGATGCA 2460  
QY 2461 GCCCAAGGGTATTAAGTTCCTGTAACCCCGTCCGCGCCACCCCTAGGTTTCGGG 2520  
DB 2461 GCCCAAGGGTATTAAGTTCCTGTAACCCCGTCCGCGCCACCCCTAGGTTTCGGG 2520  
QY 2521 GCGTATATGCTTAAGGCACATGATTCGACCTTAACATCAGAACCGGGGTAAAGCAATC 2580  
DB 2521 GCGTATATGCTTAAGGCACATGATTCGACCTTAACATCAGAACCGGGGTAAAGCAATC 2580  
QY 2581 ACCACGGGTGCCCCCATCAGTACTCCACTATGGCAAGTTTCTTGGCAGCGTGTGTC 2640  
DB 2581 ACCACGGGTGCCCCCATCAGTACTCCACTATGGCAAGTTTCTTGGCAGCGTGTGTC 2640  
QY 2641 TCTGGGGGCGCCTATGACATCAATATGATGATGAGTGCCACTCAACTGACTCGACCACT 2700

DB 2641 TCTGGGGGCGCCTATGACATCAATATGATGATGAGTGCCACTCAACTGACTCGACCACT 2700  
QY 2701 ATCTGGGCATTCGGCACAGTCTCTGGACCAACGCGAGACGCTGGAGCCGACTCTGTCGTG 2760  
DB 2701 ATCTGGGCATTCGGCACAGTCTCTGGACCAACGCGAGACGCTGGAGCCGACTCTGTCGTG 2760  
QY 2761 CTGCGCACCGCTAGCCCTCGGGATCGGTCACTGCGGCACATCCAAACATCGAGAGGTG 2820  
DB 2761 CTGCGCACCGCTAGCCCTCGGGATCGGTCACTGCGGCACATCCAAACATCGAGAGGTG 2820  
QY 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCAATCCCATCGAGACCAATC 2880  
DB 2821 GCTCTGTCCAGCACTGGAGAAATCCCTTTTATGGCAAGCAATCCCATCGAGACCAATC 2880  
QY 2881 AAGGGGGGAGGACCTCATTTCTGCCATTCACAAAGAAATGTGATGAGCTGCGCGCG 2940  
DB 2881 AAGGGGGGAGGACCTCATTTCTGCCATTCACAAAGAAATGTGATGAGCTGCGCGCG 2940  
QY 2941 AAGCTGTCCGGCCTCGGACTCAATGTGTAGCATATTAACGGGGCCTTGATGTATCCGTC 3000  
DB 2941 AAGCTGTCCGGCCTCGGACTCAATGTGTAGCATATTAACGGGGCCTTGATGTATCCGTC 3000  
QY 3001 ATACCAACTAGCGAGAGCTCATTTCTGTAGCAACGACGCTCTAATGACGGGCTTTACC 3060  
DB 3001 ATACCAACTAGCGAGAGCTCATTTCTGTAGCAACGACGCTCTAATGACGGGCTTTACC 3060  
QY 3061 GGGGATTTGCACTCAGTGTGCTGCAATACATGTGTCAACAGACAGTCACTCACTCAGC 3120  
DB 3061 GGGGATTTGCACTCAGTGTGCTGCAATACATGTGTCAACAGACAGTCACTCACTCAGC 3120  
QY 3121 CTGACCCGACCTTACCATTTGAGACGACGACGCTGCAACGAGCGGGTGTACAGTGTG 3180  
DB 3121 CTGACCCGACCTTACCATTTGAGACGACGACGCTGCAACGAGCGGGTGTACAGTGTG 3180  
QY 3181 CAGCGGCGAGGACGCTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
DB 3181 CAGCGGCGAGGACGCTGTAGGGCAGGATGGGCAATTTACAGGTTTGTGACTCCAGGA 3240  
QY 3241 GAACGGCCCTCGGGCATTTCTCGATTTCTGCGAGTGTCTATGACGGGGCTGT 3300  
DB 3241 GAACGGCCCTCGGGCATTTCTCGATTTCTGCGAGTGTCTATGACGGGGCTGT 3300  
QY 3301 GCTTGTGATCAGCTACGCGCCGCGAGACCTCAGTTAGTGTGCGGGCTTACCTAAACACA 3360  
DB 3301 GCTTGTGATCAGCTACGCGCCGCGAGACCTCAGTTAGTGTGCGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTGCGGCTCTGCGAGGACCTCTGGAGTCTGGAGAGCGTCTTTACAGGCTC 3420  
DB 3361 CCAGGTTGCGGCTCTGCGAGGACCTCTGGAGTCTGGAGAGCGTCTTTACAGGCTC 3420  
QY 3421 ACCCATATAGACCCCATTTCTTGTCCAGACTAAGCAGGACGAGACAACTTCCCTAC 3480  
DB 3421 ACCCATATAGACCCCATTTCTTGTCCAGACTAAGCAGGACGAGACAACTTCCCTAC 3480  
QY 3481 CTGGTATGATACAGGCTACGGTGTGCGGACGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
DB 3481 CTGGTATGATACAGGCTACGGTGTGCGGACGGCTCAGGCTCCACTCCATCGTGGGAC 3540  
QY 3541 CAAATGTGAAGTGTCTCATAGGCTAAAGCTTACGCTGCAAGGCGCAAGCGCCCTGCTG 3600  
DB 3541 CAAATGTGAAGTGTCTCATAGGCTAAAGCTTACGCTGCAAGGCGCAAGCGCCCTGCTG 3600  
QY 3601 TATAGCTGGGAGCCCTTCAAAACAGAGTTTACTTACACACACCCCAATACCAATACATC 3660  
DB 3601 TATAGCTGGGAGCCCTTCAAAACAGAGTTTACTTACACACACCCCAATACCAATACATC 3660  
QY 3661 ATGGCATGATGTCTGCTGACCTTGGAGGTGCTGACGAGCACTTGGGTGCTGTAGCGGA 3720  
DB 3661 ATGGCATGATGTCTGCTGACCTTGGAGGTGCTGACGAGCACTTGGGTGCTGTAGCGGA 3720  
QY 3721 GTCTTAGACGCTCTGCGCCGCTATTTGCTCTGACAAAGGAGCGTGTGCTATTTGGGCGAG 3780

Db 3721 GTCTAGCAGCTCGGCCGCTATTGCTGACAAACAGGACGCGTGGTCAATCTGCGCAGG 3780  
QY 3781 ATCATCTTTGTCGGAAAGCGGCCCATCATTTCCCGACAGGAAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTTGTCGGAAAGCGGCCCATCATTTCCCGACAGGAAAGTCTTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTGGCGCTCACACCTCCCTTATCGAACAGGAAATGACGCTCGCC 3900  
Db 3841 GATGAGATGGAAGAGTGGCGCTCACACCTCCCTTATCGAACAGGAAATGACGCTCGCC 3900  
QY 3901 GAAACAATTTCAACACAGAGCAATTCGGGTGCTGCAAAACAGCCACCAAGACGAGGCT 3960  
Db 3901 GAAACAATTTCAACACAGAGCAATTCGGGTGCTGCAAAACAGCCACCAAGACGAGGCT 3960  
QY 3961 GTGTCTCCGCTGGTGAATTCGAAGTGGCGGACCTCGAAGCCCTTCTGGCGGAAGCATATG 4020  
Db 3961 GTGTCTCCGCTGGTGAATTCGAAGTGGCGGACCTCGAAGCCCTTCTGGCGGAAGCATATG 4020  
QY 4021 TGGAAATTTTATCAGCGGGATACAATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
Db 4021 TGGAAATTTTATCAGCGGGATACAATATTTAGCAGGCTTGTCCACTCTGCTGGCAACCCC 4080  
QY 4081 GCGATAGCATCACTGATGCAATTCACAGCCTCTATCACAGCCCGGTCAACCCCAACAT 4140  
Db 4081 GCGATAGCATCACTGATGCAATTCACAGCCTCTATCACAGCCCGGTCAACCCCAACAT 4140  
QY 4141 ACCCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGTCTCTCCGAGCGT 4200  
Db 4141 ACCCTCTGTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGTCTCTCCGAGCGT 4200  
QY 4201 GCTTCTGCTTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGTGGACATAGGCCCTTGG 4260  
Db 4201 GCTTCTGCTTTTAAACATCTGGGGGATGGGTGGCGGCCCACTTGTGGACATAGGCCCTTGG 4260  
QY 4261 AAGGTGCTTTGGAATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCGCTCGTGGCC 4320  
Db 4261 AAGGTGCTTTGGAATTTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCGCTCGTGGCC 4320  
QY 4321 TTTAAGGTATGAGCGCGGAGATGCCCTCAGCGAGGACCTGGTTAACTACTTCCCTGCT 4380  
Db 4321 TTTAAGGTATGAGCGCGGAGATGCCCTCAGCGAGGACCTGGCTAACTACTTCCCTGCT 4380  
QY 4381 ATCTCTCCCTCGGCCCTAGTCGTCGGGTGCTGTGGCAGCGATATGCGTGGCAC 4440  
Db 4381 ATCTCTCCCTCGGCCCTAGTCGTCGGGTGCTGTGGCAGCGATATGCGTGGCAC 4440  
QY 4441 GTGGGCCACGGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCTTGGCTTCGCG 4500  
Db 4441 GTGGGCCACGGGAGGGGGCTGTGCAGTGGATGAACCGGCTGATAGCTTGGCTTCGCG 4500  
QY 4501 GGTAAACAGCTCTCCCGCAGCACTATGTGCTGAGAGGAGCGCTGCAGCACTGTCACT 4560  
Db 4501 GGTAAACAGCTCTCCCGCAGCACTATGTGCTGAGAGGAGCGCTGCAGCACTGTCACT 4560  
QY 4561 CAGATCCTCTAGTCTTACCATCACTCAGTCTCAAGAGGCTTCAACAGTGGATCAAC 4620  
Db 4561 CAGATCCTCTAGTCTTACCATCACTCAGTCTCAAGAGGCTTCAACAGTGGATCAAC 4620  
QY 4621 GAGGACTGTCTCAGCGCATCTCCGCTCGTGGCTAAGAGATGTTTGGGATGGATATGC 4680  
Db 4621 GAGGACTGTCTCAGCGCATCTCCGCTCGTGGCTAAGAGATGTTTGGGATGGATATGC 4680  
QY 4681 ACGGTGTGATGATTTCAAGACTGTGCTCCAGTCCAGTCTCAAGCTCTGCGGAGTTGCCGGA 4740  
Db 4681 ACGGTGTGATGATTTCAAGACTGTGCTCCAGTCCAGTCTCAAGCTCTGCGGAGTTGCCGGA 4740  
QY 4741 GTCCCTTCTTCTCATGTCAAGCTGGGTACAAGGAGTCTGGCGGGCGACGCAATCATG 4800  
Db 4741 GTCCCTTCTTCTCATGTCAAGCTGGGTACAAGGAGTCTGGCGGGCGACGCAATCATG 4800  
QY 4801 CAAACCACTGCCCCATGTGGAGCAAGATCACCGGACATGTGAAAAACGGTTCATGAGG 4860  
Db 4801 CAAACCACTGCCCCATGTGGAGCAAGATCACCGGACATGTGAAAAACGGTTCATGAGG 4860

QY 4861 ATCTGCGGCGCTAGGACCTGTAGTAACACGTGGCATGGAACATTCCTCCATTAAACGCTAC 4920  
Db 4861 ATCTGCGGCGCTAGGACCTGTAGTAACACGTGGCATGGAACATTCCTCCATTAAACGCTAC 4920  
QY 4921 ACCAGGCGCCCTGACCGCCCTCCCGCGCAAAATTTATTCAGGGCGTGTGGCGGGTG 4980  
Db 4921 ACCAGGCGCCCTGACCGCCCTCCCGCGCAAAATTTATTCAGGGCGTGTGGCGGGTG 4980  
QY 4981 GCTGTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACTAGCTGACGGGCATG 5040  
Db 4981 GCTGTGAGGAGTACGTGGAGTTACCGGGTGGGGATTTCCACTAGCTGACGGGCATG 5040  
QY 5041 ACCACTGACAAACGTAAGTGGCCCGTGTAGGTTCCGGGCCCCGAAATTTCTTCAAGAAATG 5100  
Db 5041 ACCACTGACAAACGTAAGTGGCCCGTGTAGGTTCCGGGCCCCGAAATTTCTTCAAGAAATG 5100  
QY 5101 GATGGGTGCGGTTCACAGGTACGCTCCAGCGTCAAAACCCCTCTCTACGGGAGAGTTC 5160  
Db 5101 GATGGGTGCGGTTCACAGGTACGCTCCAGCGTCAAAACCCCTCTCTACGGGAGAGTTC 5160  
QY 5161 ACATTTCTGGTGGGCTCAATCAATACCTGTGGGTTCAGCTCCCATGCGGAGCCGAA 5220  
Db 5161 ACATTTCTGGTGGGCTCAATCAATACCTGTGGGTTCAGCTCCCATGCGGAGCCGAA 5220  
QY 5221 CCGGACGTAGCAGTGTCACTTCCATGCTCACCGACCCCTCCCAATTAACGGCGGAGAG 5280  
Db 5221 CCGGACGTAGCAGTGTCACTTCCATGCTCACCGACCCCTCCCAATTAACGGCGGAGAG 5280  
QY 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTGGCCAGGTCTATCAGTAGCCAG 5340  
Db 5281 GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTCTGGCCAGGTCTATCAGTAGCCAG 5340  
QY 5341 CTGTCTGGCCTTCTTGAAGGCAACATGCACTACCCGTCACTCCCGGAGCGCTGAC 5400  
Db 5341 CTGTCTGGCCTTCTTGAAGGCAACATGCACTACCCGTCACTCCCGGAGCGCTGAC 5400  
QY 5401 CTCATCGAGGCCAACTCTGTGGGGCAGGAGATGGCGGGAAACATACCCGCGTGGAG 5460  
Db 5401 CTCATCGAGGCCAACTCTGTGGGGCAGGAGATGGCGGGAAACATACCCGCGTGGAG 5460  
QY 5461 TCAGAAAATAAGGTAGTAATTTTGAATCTTTTCGAGCCGCTCCAGCGGAGGAGTATG 5520  
Db 5461 TCAGAAAATAAGGTAGTAATTTTGAATCTTTTCGAGCCGCTCCAGCGGAGGAGTATG 5520  
QY 5521 AGGGAAGTATCCGTTCCCGCGGAGATCTTCGAGAGGTCCAGGAAATTTCCCTCGAGCGATG 5580  
Db 5521 AGGGAAGTATCCGTTCCCGCGGAGATCTTCGAGAGGTCCAGGAAATTTCCCTCGAGCGATG 5580  
QY 5581 CCCATATGGGACCGCCGGATTACAACTCTTACAGTCTTGAAGTCTCTGGAAGGACCCGGAC 5640  
Db 5581 CCCATATGGGACCGCCGGATTACAACTCTTACAGTCTTGAAGTCTCTGGAAGGACCCGGAC 5640  
QY 5641 TAGCTCCCTCCAGTGATACACGGGTTCATTTCGCGCTTCGCAAGGCCCTCCGATACCA 5700  
Db 5641 TAGCTCCCTCCAGTGATACACGGGTTCATTTCGCGCTTCGCAAGGCCCTCCGATACCA 5700  
QY 5701 CCTTCCAGGAGGAGGAGCGGTTGCTGTGAGATCTACCGTGTCTTCGCTTGGCG 5760  
Db 5701 CCTTCCAGGAGGAGGAGCGGTTGCTGTGAGATCTACCGTGTCTTCGCTTGGCG 5760  
QY 5761 GAGCTCGCACAAAGACCTTCGGCAGCTCCGAAATCTGCGCCGTTCGACAGCGGACCGCA 5820  
Db 5761 GAGCTCGCACAAAGACCTTCGGCAGCTCCGAAATCTGCGCCGTTCGACAGCGGACCGCA 5820  
QY 5821 ACGGCTCTCTTACACAGCCCTCCGACAGCGGAGTCCCGGATTCGAGCGTTCGATC 5880  
Db 5821 ACGGCTCTCTTACACAGCCCTCCGACAGCGGAGTCCCGGATTCGAGCGTTCGATC 5880  
QY 5881 TCTTCCATGCCCCCTTTCGAGGGGAGCGGGGATCCCGATCTTCAGCGAGCGGTCTTGG 5940  
Db 5881 TCTTCCATGCCCCCTTTCGAGGGGAGCGGGGATCCCGATCTTCAGCGAGCGGTCTTGG 5940



5941 TCTACCGTAAGCGAGGAGGTAGTAGGACGCTCGTCTGCTCGATGCTCTACACATGG 6000  
5941 TCTACCGTAAGCGAGGAGGTAGTAGGACGCTCGTCTGCTCGATGCTCTACACATGG 6000  
6001 ACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAACCAAGCTGCCCATCAATGCACTG 6060  
6001 ACAGCGCCCTGATCACGCCATGCGCTCGGAGGAAACCAAGCTGCCCATCAATGCACTG 6060  
6061 AGCAACTCTTTGCTCCGTCACCAACATTTGGTCTATGCTACAACTCTCGAGCGCAAGC 6120  
6061 AGCAACTCTTTGCTCCGTCACCAACATTTGGTCTATGCTACAACTCTCGAGCGCAAGC 6120  
6121 CTGCGGCAAGAGAGGTCAACCTTTGACAGACTGCGAGTCTCTGGACACCACTACCGGAC 6180  
6121 CTGCGGCAAGAGAGGTCAACCTTTGACAGACTGCGAGTCTCTGGACACCACTACCGGAC 6180  
6181 GTGCTCAAGGAGATGAAGCGAAGCGCTCCAGTTAAGGCTAAACTTTCTATCGGTGGAG 6240  
6181 GTGCTCAAGGAGATGAAGCGAAGCGCTCCAGTTAAGGCTAAACTTTCTATCGGTGGAG 6240  
6241 GAAGCTGTAGCTGACGCGCCCACTTCGCGCAGATCTAAATTTGGCTATGGGCAAG 6300  
6241 GAAGCTGTAGCTGACGCGCCCACTTCGCGCAGATCTAAATTTGGCTATGGGCAAG 6300  
6301 GACGTCCGGAACCTATCCAGCAAGCGCTTAACCACTCCGCTCGTGTGGAAGACTTG 6360  
6301 GACGTCCGGAACCTATCCAGCAAGCGCTTAACCACTCCGCTCGTGTGGAAGACTTG 6360  
6361 CTGGAAGACTGTAGACACCAATTTGACACCACTATGCGCAAAATAGGTTTCTGCG 6420  
6361 CTGGAAGACTGTAGACACCAATTTGACACCACTATGCGCAAAATAGGTTTCTGCG 6420  
6421 GTCCAAACAGAGAGGCGGCGCAAGCAGCTGCTTATGCTATGCTATGCTATGCTATG 6480  
6421 GTCCAAACAGAGAGGCGGCGCAAGCAGCTGCTTATGCTATGCTATGCTATGCTATG 6480  
6481 GTTCGTGTGTCGGAAGAAATGGCCCTTTACGATGTGCTCCACCTCCCTCAGGCGGTG 6540  
6481 GTTCGTGTGTCGGAAGAAATGGCCCTTTACGATGTGCTCCACCTCCCTCAGGCGGTG 6540  
6541 ATGGGCTCTTCAATCGGATCCCAATCTCTCTGACAGCGGCTCGAGTTCTCTGCTGAAT 6600  
6541 ATGGGCTCTTCAATCGGATCCCAATCTCTCTGACAGCGGCTCGAGTTCTCTGCTGAAT 6600  
6601 GCCTGGAAGCGAGAAATGCCCTATGGCTTCGATATGACACCGCTGTTTGACTCA 6660  
6601 GCCTGGAAGCGAGAAATGCCCTATGGCTTCGATATGACACCGCTGTTTGACTCA 6660  
6661 ACGTCACTGAGATGACATCCGCTGTTGAGGAGTCAATCTACCAATGTTGACTGTGCC 6720  
6661 ACGTCACTGAGATGACATCCGCTGTTGAGGAGTCAATCTACCAATGTTGACTGTGCC 6720  
6721 CCCGAAGCGACAGGCGCATAAAGTCTGCTACAGAGCGGCTTTATACATCGGGGCGCCCTG 6780  
6721 CCCGAAGCGACAGGCGCATAAAGTCTGCTACAGAGCGGCTTTATACATCGGGGCGCCCTG 6780  
6781 ACTAATTTCTAAAGGCGAGAACTGGGCTATCGCGTATCGCGGAGGCGGTGACTGAGC 6840  
6781 ACTAATTTCTAAAGGCGAGAACTGGGCTATCGCGTATCGCGGAGGCGGTGACTGAGC 6840  
6841 ACCAGCTGCGGTAAATACCTCTCACAATGTTACTTTGAAGCGCGCTCGGCTCTGAGCTGCG 6900  
6841 ACCAGCTGCGGTAAATACCTCTCACAATGTTACTTTGAAGCGCGCTCGGCTCTGAGCTGCG 6900  
6901 AAGCTCCAGGACTGCACGATGCTGATGCGGAGACGACCTTTGCTTATCTGTGAAAGC 6960  
6901 AAGCTCCAGGACTGCACGATGCTGATGCGGAGACGACCTTTGCTTATCTGTGAAAGC 6960  
6961 GCGGGGACCCAGAGGAGGCGGAGGCTTACGGGCTTACGAGGCTATGACTAGATAC 7020  
6961 GCGGGGACCCAGAGGAGGCGGAGGCTTACGGGCTTACGAGGCTATGACTAGATAC 7020  
7021 TCTGCCCGCTCGGGACCCGCCCAACCAAGCAATACGACTTGGAGTTGATTAACATATGC 7080

7021 TCTGCCCGCTCGGGACCCGCCCAACCAAGCAATACGACTTGGAGTTGATTAACATATGC 7080  
7081 TCCTCCAATGTGTAGTCGCGCAGCATCACTGGCAAAAGGCTGACTATCTCACCCGT 7140  
7081 TCCTCCAATGTGTAGTCGCGCAGCATCACTGGCAAAAGGCTGACTATCTCACCCGT 7140  
7141 GACCCACACCCCTTGGCGGGCTCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
7141 GACCCACACCCCTTGGCGGGCTCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
7201 TCCTGGCTAGGCAACATCATGATGCGCCCACTTGTGGGAGAGATGATCTGATG 7260  
7201 TCCTGGCTAGGCAACATCATGATGCGCCCACTTGTGGGAGAGATGATCTGATG 7260  
7261 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGAAGCCCTAGATGTCAG 7320  
7261 ACTCATTTCTTCTCCATCTTCTAGCTCAGGAACAACTTGAAGCCCTAGATGTCAG 7320  
7321 ATCTACGGGCGCTGTTACTCCATTTGAGCCACTTGAACCTTACTCAGATCAATCAAGACTC 7380  
7321 ATCTACGGGCGCTGTTACTCCATTTGAGCCACTTGAACCTTACTCAGATCAATCAAGACTC 7380  
7381 CATGSCCTTAGCGCATTTTCACTCCATGTTACTCTCCAGGTGAGATCAATAGGTTGCT 7440  
7381 CATGSCCTTAGCGCATTTTCACTCCATGTTACTCTCCAGGTGAGATCAATAGGTTGCT 7440  
7441 TCATGSCCTCAGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAGT 7500  
7441 TCATGSCCTCAGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAAGT 7500  
7501 GTCCGCGTAGGCTACTGTCTCCAGGGGGAGGCTGCCACTTGTGGCAAGTACTCTTTC 7560  
7501 GTCCGCGTAGGCTACTGTCTCCAGGGGGAGGCTGCCACTTGTGGCAAGTACTCTTTC 7560  
7561 AACTGGGCGATGAAGGACCAAGCTCAAACTCACTCCAATCCCGGCTGGTCCAGTTGAT 7620  
7561 AACTGGGCGATGAAGGACCAAGCTCAAACTCACTCCAATCCCGGCTGGTCCAGTTGAT 7620  
7621 TTATCCAGCTGTTTCTGTTGCTGTTTACAGCGGGGAGACATATATACAGCTCTCTCGT 7680  
7621 TTATCCAGCTGTTTCTGTTGCTGTTTACAGCGGGGAGACATATATACAGCTCTCTCGT 7680  
7681 GCCCGACCCGCTGGTTCATGTGTGCTACTCTCTCTGTTAGGGGTAGGCACTAT 7740  
7681 GCCCGACCCGCTGGTTCATGTGTGCTACTCTCTCTGTTAGGGGTAGGCACTAT 7740  
7741 CTACTCCCAACCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800  
7741 CTACTCCCAACCGATGAACGGGAGCTAAACACTCCAGGCCAATAGGCCATCTCTGTTT 7800  
7801 TTTCCCTT 7860  
7801 TTTCCCTT 7860  
7861 TTTTTCCTCTTTTTTCTTTTTCTTTCTTTCTTTGTTGGTCTCCATCTTAGCCCTTAGTCAGGC 7920  
7861 TTTTTCCTCTTTTTTCTTTTTCTTTCTTTCTTTGTTGGTCTCCATCTTAGCCCTTAGTCAGGC 7920  
7921 TAGCTGTGAAGTCCGTTGAGCGCTTGAAGAGAGTGTGACTGACAGAGTGTGACTGCTCTCTG 7980  
7921 TAGCTGTGAAGTCCGTTGAGCGCTTGAAGAGAGTGTGACTGACAGAGTGTGACTGCTCTCTG 7980  
7981 AGATCAAGT 7989  
7981 AGATCAAGT 7989

RESULT 15  
US-10-434-842-4  
; Sequence 4, Application US/10434842  
; Publication No. US20040005549A1  
; GENERAL INFORMATION:

; APPLICANT: Bichko, Vadim  
 ; TITLE OF INVENTION: HEPATITIS C VIRUS CONSTRUCTS CHARACTERIZED BY HIGH EFFICIENCY REPLICATION  
 ; FILE REFERENCE: 0342/1H395US3  
 ; CURRENT APPLICATION NUMBER: US 10/434,842  
 ; CURRENT FILING DATE: 2003-05-09  
 ; PRIOR APPLICATION NUMBER: US 10/233,307  
 ; PRIOR FILING DATE: 2002-08-28  
 ; PRIOR APPLICATION NUMBER: US 10/005,469  
 ; PRIOR FILING DATE: 2001-11-07  
 ; PRIOR APPLICATION NUMBER: US 60/245,866  
 ; PRIOR FILING DATE: 2000-11-07  
 ; NUMBER OF SEQ ID NOS: 17  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4  
 ; LENGTH: 7992  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: HCV9 subgenomic HCV replicon  
 ; US-10-434-842-4

Query Match 99.9%; Score 7981; DB 16; Length 7992;  
 Best Local Similarity 99.9%; Pred. No. 0;  
 Matches 7984; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY	1	GCACGCCCGATTGGGGGACACTCCACATAGATCACTCCCTGTGAGGAATCTG	60
DB	1	GCACGCCCGATTGGGGGACACTCCACATAGATCACTCCCTGTGAGGAATCTG	60
QY	61	TCTTACGAGAGAGCGTCTAGCCATGGCGTTAGTATGATGTCGACGCTCCAGGAC	120
DB	61	TCTTACGAGAGAGCGTCTAGCCATGGCGTTAGTATGATGTCGACGCTCCAGGAC	120
QY	121	CCCCCTCCGGGAGAGCCATAGTGTCTCGGAAACGGTGTACACGGAAATTCGAC	180
DB	121	CCCCCTCCGGGAGAGCCATAGTGTCTCGGAAACGGTGTACACGGAAATTCGAC	180
QY	181	GACGACCGGTCTCTTCTGGATCAACCGCTCAATGCTGAGATTTGGGCGTCCCCC	240
DB	181	GACGACCGGTCTCTTCTGGATCAACCGCTCAATGCTGAGATTTGGGCGTCCCCC	240
QY	241	GCAGAGCTGTAGCCGAGTAGTGTGGTTCGCGAAAGGCTTGTGTACTGCTGATAGG	300
DB	241	GCAGAGCTGTAGCCGAGTAGTGTGGTTCGCGAAAGGCTTGTGTACTGCTGATAGG	300
QY	301	GTGCTTGCAGAGTCCCGGAGGTCTGTAGACCGGTGACCATGACGACGAATCCTAAC	360
DB	301	GTGCTTGCAGAGTCCCGGAGGTCTGTAGACCGGTGACCATGACGACGAATCCTAAC	360
QY	361	CTCAAGAAACCAAGGGCGCCCATGATTGAACAAGATGATTGACGCGAGTTCTC	420
DB	361	CTCAAGAAACCAAGGGCGCCCATGATTGAACAAGATGATTGACGCGAGTTCTC	420
QY	421	CGCCCGCTTGGTGGAGAGCTATTGGCTATGATGGGCAACACGACAAATCGGTGCT	480
DB	421	CGCCCGCTTGGTGGAGAGCTATTGGCTATGATGGGCAACACGACAAATCGGTGCT	480
QY	481	CTGATGCGCGGTGTTCCGCTGTACGCGAGGGGCGCGGTCTTTTGTCAAGACCG	540
DB	481	CTGATGCGCGGTGTTCCGCTGTACGCGAGGGGCGCGGTCTTTTGTCAAGACCG	540
QY	541	ACCTGTCCGCTGCTCAATGAATCTGACGACGAGGCGCGGCTATCGTGGCTGGCCA	600
DB	541	ACCTGTCCGCTGCTCAATGAATCTGACGACGAGGCGCGGCTATCGTGGCTGGCCA	600
QY	601	CGACGGCGTTCCTTGGCGAGCTGTGCTGACGCTGTGCTGACGAGGAGGACTGGC	660
DB	601	CGACGGCGTTCCTTGGCGAGCTGTGCTGACGCTGTGCTGACGAGGAGGACTGGC	660
QY	661	TGCTATTGGGGAGTCCGGGGGAGGATCTCCCTGATCTACCTTGTCTTCTGCGAGA	720
DB	661	TGCTATTGGGGAGTCCGGGGGAGGATCTCCCTGATCTACCTTGTCTTCTGCGAGA	720

QY	721	AAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATTCGGCTACCTGCC	780
DB	721	AAGTATCCATCATGGCTGATGCAATGCGCGGCTGCATACGCTTGATTCGGCTACCTGCC	780
QY	781	CATTGACCAACCAAGGAAACATCGCATCGAGGAGCAGTACTCGGATGGAAGCGGTC	840
DB	781	CATTGACCAACCAAGGAAACATCGCATCGAGGAGCAGTACTCGGATGGAAGCGGTC	840
QY	841	TTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTCG	900
DB	841	TTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGCTCGCGCCAGCGAACTGTTCG	900
QY	901	CCAGGCTCAAGGCGCGCATGCGCGAGAGCATCTCGTGTGACCCATGCGCATGTCCT	960
DB	901	CCAGGCTCAAGGCGCGCATGCGCGAGAGCATCTCGTGTGACCCATGCGCATGTCCT	960
QY	961	GCTTGCAGATATCATGTGGAATGCGCGCTTTCTGATTCATCGATCTGCGCGGC	1020
DB	961	GCTTGCAGATATCATGTGGAATGCGCGCTTTCTGATTCATCGATCTGCGCGGC	1020
QY	1021	TGGGTGTGGCGACCGCTATCAGGACATAGCGTTGCTACCCGTGATATTGCTGAAGAGC	1080
DB	1021	TGGGTGTGGCGACCGCTATCAGGACATAGCGTTGCTACCCGTGATATTGCTGAAGAGC	1080
QY	1081	TTGGCGCGAATGGGCTGACCGCTTCTCGTGTTCACGATATCGCGCTCCGATTCGC	1140
DB	1081	TTGGCGCGAATGGGCTGACCGCTTCTCGTGTTCACGATATCGCGCTCCGATTCGC	1140
QY	1141	AGGGATCGCTTCTATCGCTTCTTGAGAGTCTTCTGAGTTTAAACAGACCAACAG	1200
DB	1141	AGGGATCGCTTCTATCGCTTCTTGAGAGTCTTCTGAGTTTAAACAGACCAACAG	1200
QY	1201	GTTCCTCTTAGCGGATCAATTCGCGCTTCTCCCTCCCGCTCCCGCTTACCTTACCTGC	1260
DB	1201	GTTCCTCTTAGCGGATCAATTCGCGCTTCTCCCTCCCGCTTCTCCCTCCCGCTTACCTGC	1260
QY	1261	CGAAGCGCTTGAATTAAGCGCGTGTGCTTGTCTATATGTTATTTCCACATATTG	1320
DB	1261	CGAAGCGCTTGAATTAAGCGCGTGTGCTTGTCTATATGTTATTTCCACATATTG	1320
QY	1321	CGTCTTTTGGCAATGAGGCGCGGAAACCTGCGCTTCTTCTTGAGGAGCATTCCT	1380
DB	1321	CGTCTTTTGGCAATGAGGCGCGGAAACCTGCGCTTCTTCTTGAGGAGCATTCCT	1380
QY	1381	AGGGGTCTTCTCCCTCTCGCAAGGAATGCAAGGTCTGTGTAATGTCGTGAAGAGCA	1440
DB	1381	AGGGGTCTTCTCCCTCTCGCAAGGAATGCAAGGTCTGTGTAATGTCGTGAAGAGCA	1440
QY	1441	GTTCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCTTTGCGAGCGCG	1500
DB	1441	GTTCTCTGGAAGCTTCTTGAAGACAAACAACTGTGTAGCGACCTTTGCGAGCGCG	1500
QY	1501	AACCCCGCTTGGCGAGGCTTCTGCGGCAAGGCGGCTTCTTGAGGAGCATACCT	1560
DB	1501	AACCCCGCTTGGCGAGGCTTCTGCGGCAAGGCGGCTTCTTGAGGAGCATACCT	1560
QY	1561	GCAAGCGCGCAACACCCAGTGCACGTTGTGAGTTGATGTGTGAAAGAGTCAAA	1620
DB	1561	GCAAGCGCGCAACACCCAGTGCACGTTGTGAGTTGATGTGTGAAAGAGTCAAA	1620
QY	1621	TGCTCTCTTCAAGCGTATTCAAAGGGGCTGAAGGATGCGGAGGAGTACCCCATTTG	1680
DB	1621	TGCTCTCTTCAAGCGTATTCAAAGGGGCTGAAGGATGCGGAGGAGTACCCCATTTG	1680
QY	1681	ATGGGATCTGATCTGGGCGCTCGGTGCATGCTTTACATGTGTGTAGTGTGAGTTAAAA	1740
DB	1681	ATGGGATCTGATCTGGGCGCTCGGTGCATGCTTTACATGTGTGTAGTGTGAGTTAAAA	1740
QY	1741	AACGCTTAGGCGCCCGGAAACAGCGGAGCTGTGTTTCTTTTGAAGAAACAGTAAATCC	1800
DB	1741	AACGCTTAGGCGCCCGGAAACAGCGGAGCTGTGTTTCTTTTGAAGAAACAGTAAATCC	1800
QY	1801	ATGCGCGCTATTACGGCTTACTCCCAACAGACGGGAGGCTACTTGGCTGCATCACT	1860

Db 1801 ATGGCGCTATTACGGCTACTCCCAAAGAGCGCGAGGCTACTTTGGCTGATCATCACT 1860  
QY 1861 AGCTCAGAGCGCGGAGAGAGAACAGAGTCGAGGGGAGGTCGAAAGTGGTCTCCACCGCA 1920  
Db 1861 AGCTCAGAGCGCGGAGAGAGAACAGAGTCGAGGGGAGGTCGAAAGTGGTCTCCACCGCA 1920  
QY 1921 ACACAAATCTTTCCTGGCGACCTGCGTCAATGCGGTGCTGGAGCTGTATCATGGTGCC 1980  
Db 1921 ACACAAATCTTTCCTGGCGACCTGCGTCAATGCGGTGCTGGAGCTGTATCATGGTGCC 1980  
QY 1981 GGCTCAAAAGACCCCTTTCGCGGCCCAAGGGCCCAATCACCCAAATGTACACAAATGTGGAC 2040  
Db 1981 GGCTCAAAAGACCCCTTTCGCGGCCCAAGGGCCCAATCACCCAAATGTACACAAATGTGGAC 2040  
QY 2041 CAGGACCTTCGTGGCTGGAGAGCGGCCCGCGGGCGCGTTCCTTGACACCAATGCACTGCG 2100  
Db 2041 CAGGACCTTCGTGGCTGGAGAGCGGCCCGCGGGCGCGTTCCTTGACACCAATGCACTGCG 2100  
QY 2101 GGCAGCTCGGACCTTTACTTGGTACAGAGGCATGCCGATGTCATTCGGTGGCGCGG 2160  
Db 2101 GGCAGCTCGGACCTTTACTTGGTACAGAGGCATGCCGATGTCATTCGGTGGCGCGG 2160  
QY 2161 GGCAGACAGAGGGGAGCCTACTCTCCCGCAGGCCGCTCTCTACTTGAAGGGCTCTTCG 2220  
Db 2161 GGCAGACAGAGGGGAGCCTACTCTCCCGCAGGCCGCTCTCTACTTGAAGGGCTCTTCG 2220  
QY 2221 GGCAGTCCAGCTGCTGCGCTCGGGGACAGCGTGTGGGCACTTTTCGGGCTGCGCGTGGC 2280  
Db 2221 GGCAGTCCAGCTGCTGCGCTCGGGGACAGCGTGTGGGCACTTTTCGGGCTGCGCGTGGC 2280  
QY 2281 ACCGAGGGGTTGGAGGGGCTGACCTTGTACCCGTCGAGTCTATGGGAACCACTATG 2340  
Db 2281 ACCGAGGGGTTGGAGGGGCTGACCTTGTACCCGTCGAGTCTATGGGAACCACTATG 2340  
QY 2341 CGGTCCCGGCTTTCACGGCAAACTCGTCCGCTCCGGCCGTACCGCAGACATTCAGGTG 2400  
Db 2341 CGGTCCCGGCTTTCACGGCAAACTCGTCCGCTCCGGCCGTACCGCAGACATTCAGGTG 2400  
QY 2401 GCCCATCTACAGCCCTTACTGTTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2460  
Db 2401 GCCCATCTACAGCCCTTACTGTTAGCGGCAAGAGCACTAAGGTGCGGCTGCGTATGCA 2460  
QY 2461 GCCCAAGGATTAAGGTGCTTCTCTGAAACCGGTCGCGGCAACCCCTAGGTTCGGG 2520  
Db 2461 GCCCAAGGATTAAGGTGCTTCTCTGAAACCGGTCGCGGCAACCCCTAGGTTCGGG 2520  
QY 2521 GCGTATATGCTAAGGCACATGTTATCGACCCCTAACATCAGAACCGGGGTAAGGACCATC 2580  
Db 2521 GCGTATATGCTAAGGCACATGTTATCGACCCCTAACATCAGAACCGGGGTAAGGACCATC 2580  
QY 2581 ACACGGGTGCCCCCATCAGCTACTCCACTATGGCAAGTTCTTTGCGGACGGTGGTTC 2640  
Db 2581 ACACGGGTGCCCCCATCAGCTACTCCACTATGGCAAGTTCTTTGCGGACGGTGGTTC 2640  
QY 2641 TCTGGGGCGGCTATGATCATCAATATGATGATGAGTGCCACTCACTGACTCGACCACT 2700  
Db 2641 TCTGGGGCGGCTATGATCATCAATATGATGATGAGTGCCACTCACTGACTCGACCACT 2700  
QY 2701 ATCTGGGCACTCGGCACAGTCTTGGAACCAAGCGGAGACGGCTGGAGCGGCACTCGTGTG 2760  
Db 2701 ATCTGGGCACTCGGCACAGTCTTGGAACCAAGCGGAGACGGCTGGAGCGGCACTCGTGTG 2760  
QY 2761 CTCCGCAACCGCTAGCCTTCGGGATTCGGTCAACCTGCGACATCCAAACATCGAGAGGTG 2820  
Db 2761 CTCCGCAACCGCTAGCCTTCGGGATTCGGTCAACCTGCGACATCCAAACATCGAGAGGTG 2820  
QY 2821 GCTCTGTCAGCACTGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
Db 2821 GCTCTGTCAGCACTGAGAAATCCCTTTTATGGCAAGCCATCCCATCGAGACCATC 2880  
QY 2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCACAGAAAGAAATGTGATGATGATGCTCGCGG 2940

Db 2881 AAGGGGGGAGGACCTCATTTTCTGCCATTCACAGAAAGAAATGTGATGAGTCCCGGG 2940  
QY 2941 AAGTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTTGATGTATCCCGTC 3000  
Db 2941 AAGTGTCCGGCTCGGACTCAATGCTGTAGCATATTACCGGGGCTTTGATGTATCCCGTC 3000  
QY 3001 ATACCACTAGCGGAGAGCTCATTTGTGCTAGCAACGAGCGCTTAATGACGGGCTTTACC 3060  
Db 3001 ATACCACTAGCGGAGAGCTCATTTGTGCTAGCAACGAGCGCTTAATGACGGGCTTTACC 3060  
QY 3061 GGCATTTCCGACTCAGTGTATGATCGACTCAATATGTCACCCAGACAGTGCAGTTCAGC 3120  
Db 3061 GGCATTTCCGACTCAGTGTATGATCGACTCAATATGTCACCCAGACAGTGCAGTTCAGC 3120  
QY 3121 CTGACCCGCACTTCACCAATTGAGACGACCGGTGCGCAAGACGCGGTGTCAAGCTCG 3180  
Db 3121 CTGACCCGCACTTCACCAATTGAGACGACCGGTGCGCAAGACGCGGTGTCAAGCTCG 3180  
QY 3181 CAGCGGCGAGGACAGACTGCTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGCA 3240  
Db 3181 CAGCGGCGAGGACAGACTGCTAGGGGAGGATGGGCATTTACAGGTTTGTGACTCCAGCA 3240  
QY 3241 GAAAGGCGCTCGGCACTGTTCCGATTCCTCGGTTCTGTCCGAGTGTATGACGCGGCTGT 3300  
Db 3241 GAAAGGCGCTCGGCACTGTTCCGATTCCTCGGTTCTGTCCGAGTGTATGACGCGGCTGT 3300  
QY 3301 GCTTGTACGAGCTCAGCGCCGCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
Db 3301 GCTTGTACGAGCTCAGCGCCGCGAGACCTCAGTTAGTTGCGGGCTTACCTAAACACA 3360  
QY 3361 CCAGGTTGCCCTGCTGCCAGGACCATCTGGAGTTCTGGAGAGCGTCTTTACAGGCGTC 3420  
Db 3361 CCAGGTTGCCCTGCTGCCAGGACCATCTGGAGTTCTGGAGAGCGTCTTTACAGGCGTC 3420  
QY 3421 ACCCACATAGACGCCCTTCTTCCAGACTAAGCAGCAGAGACAACTCCCGCTAC 3480  
Db 3421 ACCCACATAGACGCCCTTCTTCCAGACTAAGCAGCAGAGACAACTCCCGCTAC 3480  
QY 3481 CTGCTAGCATACCAAGGCTACGGTGTGGCCAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
Db 3481 CTGCTAGCATACCAAGGCTACGGTGTGGCCAGGCTCAGGCTCCACCTCCATCGTGGAC 3540  
QY 3541 CAATGTGGAAGTGTCTCATACGCTAAAGCTAACGCTACGCTGACGCGGCCAACCGCCCTGCTG 3600  
Db 3541 CAATGTGGAAGTGTCTCATACGCTAAAGCTAACGCTACGCTGACGCGGCCAACCGCCCTGCTG 3600  
QY 3601 TATAGGCTGGAGCGGTTTCAAAACGAGTTACTACACACAGCCCAATACCAATACATC 3660  
Db 3601 TATAGGCTGGAGCGGTTTCAAAACGAGTTACTACACACAGCCCAATACCAATACATC 3660  
QY 3661 ATGGCATGCTGCTGGCTGACCTGGAGTGTCTACAGAGCACTGGGTGCTGGTAGGGCGGA 3720  
Db 3661 ATGGCATGCTGCTGGCTGACCTGGAGTGTCTACAGAGCACTGGGTGCTGGTAGGGCGGA 3720  
QY 3721 GTCTAGAGCTGTGGCGGCTATTGCTGACAAACAGGAGCGTGTGATTTGGGGAGG 3780  
Db 3721 GTCTAGAGCTGTGGCGGCTATTGCTGACAAACAGGAGCGTGTGATTTGGGGAGG 3780  
QY 3781 ATCATCTTGTCCGAAAGCGGCTCATTTCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
Db 3781 ATCATCTTGTCCGAAAGCGGCTCATTTCCGACAGGGAAGTCTTTTACCGGAGTTC 3840  
QY 3841 GATGAGATGGAAGAGTGCCTCACCTCTTACATCGAAACAGGGAATGCAAGTCCGC 3900  
Db 3841 GATGAGATGGAAGAGTGCCTCACCTCTTACATCGAAACAGGGAATGCAAGTCCGC 3900  
QY 3901 GAACATTTAAACAGAAAGCAATCGGGTGTCTGGAACAGCCCAAGAGCGAGGCT 3960  
Db 3901 GAACATTTAAACAGAAAGCAATCGGGTGTCTGGAACAGCCCAAGAGCGAGGCT 3960  
QY 3961 GCTGCTCCCGTGTGGAATCCAGTGCAGACCTCGAAGCTTCTGGGCGAAGCATATG 4020  
Db 3961 GCTGCTCCCGTGTGGAATCCAGTGCAGACCTCGAAGCTTCTGGGCGAAGCATATG 4020

QY	4021	TGGAATTTTCATCAGCGGGATACAAATATTTAGCAGGCTTGTCCAATCTGCTGSCAACCC	4080	QY	5101	GATGGGTCGGTTCAGCAGGTAACGCTCCAGCGTGCAAAACCCCTCTCTAAGGAGGAGGTC	5160
Db	4021	TGGAATTTTCATCAGCGGGATACAAATATTTAGCAGGCTTGTCCAATCTGCTGSCAACCC	4080	Db	5101	GATGGGTCGGTTCAGCAGGTAACGCTCCAGCGTGCAAAACCCCTCTCTAAGGAGGAGGTC	5160
QY	4081	GGGATAGCATCACTGATGGCAATTCACAGCCTCTATCACCAGCGCGCTCACCAACCAAT	4140	QY	5161	ACATTCCTGCTGGTTCGGGTCAATCAATACCTGGTGGGTACAGCTCCATTCGAGCCCGAA	5220
Db	4081	GGGATAGCATCACTGATGGCAATTCACAGCCTCTATCACCAGCGCGCTCACCAACCAAT	4140	Db	5161	ACATTCCTGCTGGTTCGGGTCAATCAATACCTGGTGGGTACAGCTCCATTCGAGCCCGAA	5220
QY	4141	ACCTCCTGTTTAAATCTCTGGGGGATGGGTGGCGCCCACTTGTCTCTCCAGCGCT	4200	QY	5221	CGGAGCTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTCAGCGGAGACG	5280
Db	4141	ACCTCCTGTTTAAATCTCTGGGGGATGGGTGGCGCCCACTTGTCTCTCCAGCGCT	4200	Db	5221	CGGAGCTAGCAGTGTCTCACTTCCATGCTCACCGACCCCTCCACATTCAGCGGAGACG	5280
QY	4201	GCTTCTGCTTTTACATCTGAGCGCGCATCTGAGCGCGCTTGTGGCAGCATAGAGCCCTGGG	4260	QY	5281	GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTGCGAGCTCAATCAGCTAGCCAG	5340
Db	4201	GCTTCTGCTTTTACATCTGAGCGCGCATCTGAGCGCGCTTGTGGCAGCATAGAGCCCTGGG	4260	Db	5281	GCTAAGCGTAGGCTGGCCAGGGGATCTCCCGCTCTGCGAGCTCAATCAGCTAGCCAG	5340
QY	4261	AAGGTCTGTGGATATTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCGCTCTGGGC	4320	QY	5341	CTGTCTGCGCCTTCTTGAAGGCAATGCACTACCCGCTCATGACTCCCGGACGCTGAC	5400
Db	4261	AAGGTCTGTGGATATTTGGCAGGTTATGGAGCAGGGGTGGCAGCGCGCTCTGGGC	4320	Db	5341	CTGTCTGCGCCTTCTTGAAGGCAATGCACTACCCGCTCATGACTCCCGGACGCTGAC	5400
QY	4321	TTTAAAGTCATGAGCGCGAGATGCCCTCCACGAGGACCTGTTAACTACTCCCTGCT	4380	QY	5401	CTCATCGAGCCAAACCTCTGCTGGCGGAGGATGGCGGGAACATCACCCGCTGGAG	5460
Db	4321	TTTAAAGTCATGAGCGCGAGATGCCCTCCACGAGGACCTGTTAACTACTCCCTGCT	4380	Db	5401	CTCATCGAGCCAAACCTCTGCTGGCGGAGGATGGCGGGAACATCACCCGCTGGAG	5460
QY	4381	ATCCTCTCCCTGCGCCCTAGTCTGCGGGTCTGTGCGCAGCGATACCTGCGTCCGCGAC	4440	QY	5461	TCAGAAAAAAGGTAGTAAATTTTGGACTCTTTTGGAGCTCTTTCAGCGCGCTCCAGCGGAGGATGAG	5520
Db	4381	ATCCTCTCCCTGCGCCCTAGTCTGCGGGTCTGTGCGCAGCGATACCTGCGTCCGCGAC	4440	Db	5461	TCAGAAAAAAGGTAGTAAATTTTGGAGCTCTTTCAGCGCGCTCCAGCGGAGGATGAG	5520
QY	4441	GTGGGCCCCAGGGGAGGGGCTGTGACGTGGATGAACCGGCTGATAGCGTTTCGCTTCGCGG	4500	QY	5521	AGGAAATATCCTTTCGCGGAGATCTCGGAGGTCGAGGAAATTCCTCCAGCGATG	5580
Db	4441	GTGGGCCCCAGGGGAGGGGCTGTGACGTGGATGAACCGGCTGATAGCGTTTCGCTTCGCGG	4500	Db	5521	AGGAAATATCCTTTCGCGGAGATCTCGGAGGTCGAGGAAATTCCTCCAGCGATG	5580
QY	4501	GGTAAACAGCTCTCCCGCCAGCAGTATGCTCTGAGAGGACGCTGAGCAGCTGTCAC	4560	QY	5581	CCCATATGGCAGCGCCGATTAACAACCTCTCACTGTTAGAGTCTCTGGAAGGACCCGAC	5640
Db	4501	GGTAAACAGCTCTCCCGCCAGCAGTATGCTCTGAGAGGACGCTGAGCAGCTGTCAC	4560	Db	5581	CCCATATGGCAGCGCCGATTAACAACCTCTCACTGTTAGAGTCTCTGGAAGGACCCGAC	5640
QY	4561	CAGATCCTCTAGTCTTACCATCACTCAGCTCTGTAAGAGGCTTCAACAGTGGATCAAC	4620	QY	5641	TACGTCCCTCCAGTGTACAGCGGTGTCCATTCGCCCTGCTCAAGGCCCTCCGATACCA	5700
Db	4561	CAGATCCTCTAGTCTTACCATCACTCAGCTCTGTAAGAGGCTTCAACAGTGGATCAAC	4620	Db	5641	TACGTCCCTCCAGTGTACAGCGGTGTCCATTCGCCCTGCTCAAGGCCCTCCGATACCA	5700
QY	4621	GAGGACTGTCTCAGCGCATCTCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC	4680	QY	5701	CCTCAGGAGGAGGAGGAGCGTTGCTCTGTCAGATCTACCGTGTCTTCTGCTTGGCGG	5760
Db	4621	GAGGACTGTCTCAGCGCATCTCGGCTCGTGGCTAAGAGATGTTTGGGATTTGGATATGC	4680	Db	5701	CCTCAGGAGGAGGAGGAGCGTTGCTCTGTCAGATCTACCGTGTCTTCTGCTTGGCGG	5760
QY	4681	ACGGTGTGACTGATTTCAAGACTCTGAGTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG	4740	QY	5761	GAGCTCGCACAAAGACCTTCGCGAGCTCCGATTCGCGCGCTCGACAGCGGACGCGCA	5820
Db	4681	ACGGTGTGACTGATTTCAAGACTCTGAGTCTGAGTCTGAGTCTGAGTCTGAG	4740	Db	5761	GAGCTCGCACAAAGACCTTCGCGAGCTCCGATTCGCGCGCTCGACAGCGGACGCGCA	5820
QY	4741	GTCCCTCTCTCATGTCAAAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAG	4800	QY	5821	ACGGCCTCTCTGACAGCGCTCCGACGAGCGGAGTCCGAGCTGAGTCTGAGTCTGAG	5880
Db	4741	GTCCCTCTCTCATGTCAAAGGAGTCTGAGGAGTCTGAGGAGTCTGAGGAGTCTGAG	4800	Db	5821	ACGGCCTCTCTGACAGCGCTCCGACGAGCGGAGTCCGAGCTGAGTCTGAGTCTGAG	5880
QY	4801	CAAAACACCTGCGCATGTGAGCAGACATCAGCGGACATGTGAAAAACGTTTCCATGAGG	4860	QY	5881	TCCTCATGCCCCCTTGGGGGAGCGGGGATCCCGATCTCAGCGAGCGGTCTTGG	5940
Db	4801	CAAAACACCTGCGCATGTGAGCAGACATCAGCGGACATGTGAAAAACGTTTCCATGAGG	4860	Db	5881	TCCTCATGCCCCCTTGGGGGAGCGGGGATCCCGATCTCAGCGAGCGGTCTTGG	5940
QY	4861	ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGATGGAACATTTCCCCATTAACGGGTAC	4920	QY	5941	TCTACCGTAAAGCAGGAGGCTAGTGAAGGACCTGCTGCTGCTGCTGCTGCTTACATAGG	6000
Db	4861	ATCGTGGGCTTAGGACCTGTAGTAAACAGTGGATGGAACATTTCCCCATTAACGGGTAC	4920	Db	5941	TCTACCGTAAAGCAGGAGGCTAGTGAAGGACCTGCTGCTGCTGCTGCTTACATAGG	6000
QY	4921	ACCACGGGCTTGCAGCGCTTCCCGGCGCCAAATTTATTTCTAGGGGCTGTGGCGGCTG	4980	QY	6001	ACAGGCGCCTGATCAGCGCATCGGTGCGGAGGAAACCAAGCTGCGCATCAATGCACTG	6060
Db	4921	ACCACGGGCTTGCAGCGCTTCCCGGCGCCAAATTTATTTCTAGGGGCTGTGGCGGCTG	4980	Db	6001	ACAGGCGCCTGATCAGCGCATCGGTGCGGAGGAAACCAAGCTGCGCATCAATGCACTG	6060
QY	4981	GCTGCTGAGGAGTACGTGGAGTGTACCGGGGTGGGGATTTTCCACTACGTGACGGGCATG	5040	QY	6061	AGCAACTCTTGTCTCGTCAACACCTTGGTCTATGCTAACAATCTCGAGCGGCAAGC	6120
Db	4981	GCTGCTGAGGAGTACGTGGAGTGTACCGGGGTGGGGATTTTCCACTACGTGACGGGCATG	5040	Db	6061	AGCAACTCTTGTCTCGTCAACACCTTGGTCTATGCTAACAATCTCGAGCGGCAAGC	6120
QY	5041	ACCACTGACAAACGTAAGTGCCTGTCAGTTCGCGCCCGGAAATTTCTTACAGAAAGTG	5100	QY	6121	CTCGGCGAGAAAGGTCACCTTTTGACAGACTGCGAGTCTCTGGAAGGACCTACCGGAC	6180
Db	5041	ACCACTGACAAACGTAAGTGCCTGTCAGTTCGCGCCCGGAAATTTCTTACAGAAAGTG	5100	Db	6121	CTCGGCGAGAAAGGTCACCTTTTGACAGACTGCGAGTCTCTGGAAGGACCTACCGGAC	6180
				QY	6181	GTGCTCAAGGAGATGAAGCGGAGGCGTCCACAGTTAAGGCTAAATTTCTATCCGCTGAG	6240

Db 6181 GTGCTCAAGGAGATGAAGGCGAGGCTCCACAGTTAAGGCTAAACTCTATCCGTTGAG 6240  
QY 6241 GAAGCTGTAGCTGAGCGCCCAATGCGGCCAGATCTAAATTTGGCTATGGGCCAAAG 6300  
Db 6241 GAAGCTGTAGCTGAGCGCCCAATGCGGCCAGATCTAAATTTGGCTATGGGCCAAAG 6300  
QY 6301 GACGTCCGGAACCTATCAGGAAGCGCTTAACACATCCGCTCCGTTGGAGGACTTG 6360  
Db 6301 GACGTCCGGAACCTATCAGGAAGCGCTTAACACATCCGCTCCGTTGGAGGACTTG 6360  
QY 6361 CTGGAAGACACTGACACACCAATTCACACCAATCATGCGCAAAAAATGAGGTTTCTGC 6420  
Db 6361 CTGGAAGACACTGACACACCAATTCACACCAATCATGCGCAAAAAATGAGGTTTCTGC 6420  
QY 6421 GTCCAAACAGAGAAAGGGGGCGCAAGCGAGCTCGCTTATCGTATTCACAGATTGGG 6480  
Db 6421 GTCCAAACAGAGAAAGGGGGCGCAAGCGAGCTCGCTTATCGTATTCACAGATTGGG 6480  
QY 6481 GTTCGTGTGCGGAAATGSCCTTTACGATGTGTTCCACCTCCCTCAGGCCGTG 6540  
Db 6481 GTTCGTGTGCGGAAATGSCCTTTACGATGTGTTCCACCTCCCTCAGGCCGTG 6540  
QY 6541 ATGGCTCTTTCATACGGAATTCGAATCTCTCTGACAGCGGTTCGAGTTCTGTTGAAT 6600  
Db 6541 ATGGCTCTTTCATACGGAATTCGAATCTCTCTGACAGCGGTTCGAGTTCTGTTGAAT 6600  
QY 6601 GCCTGGAAGCGAAGAAATGSCCTATGGGCTTCGCAATGACACCCGCTGTTTGAATCA 6660  
Db 6601 GCCTGGAAGCGAAGAAATGSCCTATGGGCTTCGCAATGACACCCGCTGTTTGAATCA 6660  
QY 6661 ACAGTCACTGAGATGACATCGCTGTTGAGGAGTCAATCTACCAATGTTGACTTGCC 6720  
Db 6661 ACAGTCACTGAGATGACATCGCTGTTGAGGAGTCAATCTACCAATGTTGACTTGCC 6720  
QY 6721 CCCGAAGCGACAGACGCAATGAGTCTGCTCAGAGCGCTTTACATCGGGGGCCCCCTG 6780  
Db 6721 CCCGAAGCGACAGACGCAATGAGTCTGCTCAGAGCGCTTTACATCGGGGGCCCCCTG 6780  
QY 6781 ACTAATTTAAAGGCGAAGCTGCGGCTATCGCGGTGCGCGGAGCGGTGTAATGACG 6840  
Db 6781 ACTAATTTAAAGGCGAAGCTGCGGCTATCGCGGTGCGCGGAGCGGTGTAATGACG 6840  
QY 6841 ACCAGTCCGCTGATACCTTACATGTTACTTGAAGCGCGCTGCGGCTGCTGAGCTGCG 6900  
Db 6841 ACCAGTCCGCTGATACCTTACATGTTACTTGAAGCGCGCTGCGGCTGCTGAGCTGCG 6900  
QY 6901 AAGCTCCAGGACTGCACGATGCTCGTATGCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960  
Db 6901 AAGCTCCAGGACTGCACGATGCTCGTATGCGGAGACGACCTTGTCTGTTATCTGTGAAAGC 6960  
QY 6961 GCGGGGACCCAAAGAGGACGAGCGAGCGCTACGCGGCTTCAAGGCTTATGACTAGATAC 7020  
Db 6961 GCGGGGACCCAAAGAGGACGAGCGAGCGCTACGCGGCTTCAAGGCTTATGACTAGATAC 7020  
QY 7021 TCTGCCCCCTGGGAGCGCGCCAAACAGAAATACGACTTGGAGTTGATTAACATATGC 7080  
Db 7021 TCTGCCCCCTGGGAGCGCGCCAAACAGAAATACGACTTGGAGTTGATTAACATATGC 7080  
QY 7081 TCCTTCCAAATGTGTCAGTCCGCGACGATGATCTGGCAAAAGGTTGATCTATCTACCCGT 7140  
Db 7081 TCCTTCCAAATGTGTCAGTCCGCGACGATGATCTGGCAAAAGGTTGATCTATCTACCCGT 7140  
QY 7141 GACCCCAACACCCCTTTGCGGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
Db 7141 GACCCCAACACCCCTTTGCGGGGCTGCGTGGGAGACAGCTAGACACACTCCAGTCAAT 7200  
QY 7201 TCCTGGCTAGGCAACATCATGATGATGGCCACCTTGTGGGCAAGGATGATCTGATG 7260  
Db 7201 TCCTGGCTAGGCAACATCATGATGATGGCCACCTTGTGGGCAAGGATGATCTGATG 7260  
QY 7261 ACTCATTTTCTCCATCCTTCTAGCTCAGGAACAATTTGAAAGCGCTAGATTGTCTAG 7320

Db 7261 ACTCATTTTCTCCATCCTTCTAGCTCAGGAACAACCTTGAAAGAGCCCTAGATTGTCTAG 7320  
QY 7321 ATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTACCTCAGATCATTTCAACGACTC 7380  
Db 7321 ATCTACGGGGCTGTACTCCATTGAGCCACTTGACCTACCTCAGATCATTTCAACGACTC 7380  
QY 7381 CATGGCCTTAGGCAATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGT 7440  
Db 7381 CATGGCCTTAGGCAATTTTCACTCCATAGTTACTCTCCAGGTGAGATCAATAGGTTGGT 7440  
QY 7441 TCATGCCCTCAGGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAACT 7500  
Db 7441 TCATGCCCTCAGGAAACTTGGGGTACCGCCCTTGGAGTCTGGAGACATCGGGCCAGAACT 7500  
QY 7501 GTCCGCGCTAGGCTACTCTCCAGGGGGGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
Db 7501 GTCCGCGCTAGGCTACTCTCCAGGGGGGGGCTGCCACTTGTGGCAAGTACCTCTTC 7560  
QY 7561 AACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAATCCCGGCTGCGTCCAGTTGGAT 7620  
Db 7561 AACTGGGCGAGTAAGGACCAAGCTCAAACTCACTCCAATCCCGGCTGCGTCCAGTTGGAT 7620  
QY 7621 TTATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAGAGCTGTCTCGT 7680  
Db 7621 TTATCCAGCTGGTTCGTTGCTGGTTACAGCGGGGAGACATATATCAGAGCTGTCTCGT 7680  
QY 7681 GCGGACCCCGCTGCTCATGTGGTGCCTACTCCTACTTCTGTAGGGGTAGGCATCTAT 7740  
Db 7681 GCGGACCCCGCTGCTCATGTGGTGCCTACTCCTACTTCTGTAGGGGTAGGCATCTAT 7740  
QY 7741 CTACTCCCAACCCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800  
Db 7741 CTACTCCCAACCCGATGAACGGGACCTAAACACTCCAGGCCAATAGGCCATCTCTTTT 7800  
QY 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
Db 7801 TTTCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7860  
QY 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
Db 7861 TTTTTCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 7920  
QY 7921 TAGCTGTGAAAGGTCCTGAGCGGCTGAGCTGAGAGAGTCTGATATCTGGCTCTCTGC 7980  
Db 7921 TAGCTGTGAAAGGTCCTGAGCGGCTGAGCTGAGAGAGTCTGATATCTGGCTCTCTGC 7980  
QY 7981 AGATCAAGT 7989  
Db 7981 AGATCAAGT 7989

Search completed: December 9, 2004, 13:40:33  
Job time : 3611 secs

**This Page Blank (uspto)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 8, 2004, 20:11:15 ; Search time 21853 Seconds

(without alignments)  
13321.604 Million cell updates/sec

Title: US-09-576-989-6

Perfect score: 7989

Sequence: 1 gccagcccccggatggggc.....ggcctctctcagatcaagt 7989

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.\*

1: gb\_est1.\*

2: gb\_est2.\*

3: gb\_est3.\*

4: gb\_est4.\*

5: gb\_est5.\*

6: gb\_est6.\*

7: gb\_est7.\*

8: gb\_est8.\*

9: gb\_est9.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	795.4	10.0	935	7	CK284786 EST747508
2	794	9.9	804	7	CK291519 EST754233
3	794	9.9	856	7	CK287297 EST750019
4	794	9.9	910	7	CK287930 EST750652
5	794	9.9	933	7	CK291799 EST754513
6	794	9.9	936	7	CK256977 EST740614
7	794	9.9	947	7	CK298208 EST760922
8	794	9.9	954	7	CK283361 EST746083
9	778.8	9.7	811	7	CK288185 EST750907
10	587	7.3	878	7	CK288711 EST751433
11	577.4	7.2	789	8	AQ361914 mgxb0005K
12	559.8	7.0	620	8	AQ398387 mgxb0009J
13	542.4	6.8	549	7	CK801630 NF32c08f4
14	514.8	6.4	561	8	AQ447874 mgxb0012I
15	508	6.4	509	8	AQ447775 mgxb0011E
16	505.4	6.3	637	8	AQ447230 mgxb0006O
17	503.8	6.3	592	8	AQ449162 mgxb0023E
18	499	6.2	499	8	AQ398880 mgxb0006P
19	496.4	6.2	498	8	AQ398730 mgxb0009G
20	477.4	6.0	586	7	CK801096 NF15h06f4
21	474.8	6.0	857	8	AQ875013 V120B8 mt
22	474.8	5.9	710	8	BH972732 odj46h10
23	469.8	5.9	473	8	AQ397768 mgxb0001J
24	468.8	5.9	594	8	AQ397253 mgxb0003C

C 25	465.2	5.8	850	8	AQ875052	AQ875052 V120F5 mt
C 26	463.4	5.8	798	8	AQ876139	AQ876139 V133G10 m
C 27	458.8	5.7	482	1	AL449872	AL449872 AL449872
C 28	456.8	5.7	861	8	AQ875040	AQ875040 V120B4 mt
C 29	455.6	5.7	815	8	AQ876134	AQ876134 V133F5 mt
C 30	451.2	5.6	799	8	AQ876220	AQ876220 V152G7 mt
C 31	450	5.6	784	8	AQ876152	AQ876152 V133H12 m
C 32	444.6	5.6	790	8	AQ875904	AQ875904 V130H11 m
C 33	437.2	5.5	797	8	AQ876297	AQ876297 V153G3 mt
C 34	434.6	5.4	765	7	CK283362	CK283362 EST746084
C 35	426.8	5.3	804	8	AQ876080	AQ876080 V133A4 mt
C 36	422.8	5.3	759	7	CK298209	CK298209 EST760923
C 37	414.6	5.2	791	8	AQ876019	CK298209 V132C3 mt
C 38	409.4	5.1	748	8	AQ501530	AQ501530 V20F6 mt
C 39	408	5.1	635	8	AQ447643	AQ447643 mgxb0008M
C 40	408	5.1	656	8	AQ447140	AQ447140 mgxb0005I
C 41	404	5.1	790	8	AQ875912	AQ875912 V130H8 mt
C 42	404	5.1	791	8	AQ876201	AQ876201 V152E2 mt
C 43	402.6	5.0	856	8	AQ875050	AQ875050 V120F2 mt
C 44	401.8	5.0	760	7	CK287931	CK287931 EST750653
C 45	398.8	5.0	722	6	CD641598	CD641598 AGENCOURT

## ALIGNMENTS

RESULT 1  
CK284786  
LOCUS EST747508 Nicotiana benthamiana mixed tissue cdNA library, linear EST 02-AUG-2004  
DEFINITION normalised, full-length Nicotiana benthamiana cDNA clone NEMAQ41 5', end, mRNA sequence.  
ACCESSION CK284786  
VERSION CK284786.1 GI:39858698  
KEYWORDS EST.  
SOURCE Nicotiana benthamiana  
ORGANISM Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Nicotiana.  
REFERENCE 1 (bases 1 to 935)  
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B., Skaskawicz, B., Jin, H. and Baker, B.  
TITLE Generation of EST sequences from Nicotiana benthamiana  
JOURNAL Unpublished (2003)  
COMMENT Other ESTs: EST747509  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

## FEATURES

Location/Qualifiers  
1..935  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NEMAQ41"  
/tissue\_type="abiotic and biotic stress-treated leaves, callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA library, normalized, full-length"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from

these tissues and pooled in approximately equal molar amounts."

ORIGIN	Query Match	10.0%;	Score 795.4;	DB 7;	Length 935;
	Best Local Similarity	99.9%;	Pred. No. 3.6e-174;		
	Matches 796;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY	386	CATGATGAAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATT	445		
DB	31	CATGATGAAACAAGATGGATTGCACGCAGGTTCTCCGGCCGCTTGGGTGGAGAGGCTATT	90		
QY	446	CGGCTATGACTGGGCAACAAGACAATCGGCTGCTCTGATGCCGCCGTGTTCGGGCTGTC	505		
DB	91	CGGCTATGACTGGGCAACAAGACAATCGGCTGCTCTGATGCCGCCGTGTTCGGGCTGTC	150		
QY	506	AGCGCAGGGCGCCCGGTTCTTTTGTCAAGACGACCTGTCCGGTGCCCTGAATGAAC	565		
DB	151	AGCGCAGGGCGCCCGGTTCTTTTGTCAAGACGACCTGTCCGGTGCCCTGAATGAAC	210		
QY	566	GCAGGACGAGCAGCGGGCTATCTGGCTGGCCACGACGCGGGCTTCTTTGGCGAGCTGT	625		
DB	211	GCAGGACGAGCAGCGGGCTATCTGGCTGGCCACGACGCGGGCTTCTTTGGCGAGCTGT	270		
QY	626	GCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGTCTATTGGGCGAAAGTGCCGGGGCA	685		
DB	271	GCTCGACGTTGTCACTGAAGCGGGAAGGACTGGCTGTCTATTGGGCGAAAGTGCCGGGGCA	330		
QY	686	GGATCTCTGTCTATCTACCTTGTCTCTCGGAGAAAGTATCCATGATGCTCATGCAAT	745		
DB	331	GGATCTCTGTCTATCTACCTTGTCTCTCGGAGAAAGTATCCATGATGCTCATGCAAT	390		
QY	746	CGCGCGGCTGCATAACGTTGATCCGGTACTCTGCCATTCGACACCAACGCAAAACATCG	805		
DB	391	CGCGCGGCTGCATAACGTTGATCCGGTACTCTGCCATTCGACACCAACGCAAAACATCG	450		
QY	806	CATCGACGAGCAGCTACTCGGATGAGCGGCTCTTGTCGATCAGGATGATCTGACGGA	865		
DB	451	CATCGACGAGCAGCTACTCGGATGAGCGGCTCTTGTCGATCAGGATGATCTGACGGA	510		
QY	866	AGAGCATCAGGGGCTCGCGCAGCCGAATGTTTCGCCAGGCTCAAGCGCGCATGCCCGA	925		
DB	511	AGAGCATCAGGGGCTCGCGCAGCCGAATGTTTCGCCAGGCTCAAGCGCGCATGCCCGA	570		
QY	926	CGGCGAGGATCTCGTGTGTGACCATGCGGATGCTTGCTTCGCGAAATATCATGTGTGAAAA	985		
DB	571	CGGCGAGGATCTCGTGTGTGACCATGCGGATGCTTGCTTCGCGAAATATCATGTGTGAAAA	630		
QY	986	TGGCCGCTTTCTGGATTTCATCGACTGTGCCGGCTGGGTGGCGAACCGCTATCAGGA	1045		
DB	631	TGGCCGCTTTCTGGATTTCATCGACTGTGCCGGCTGGGTGGCGAACCGCTATCAGGA	690		
QY	1046	CATAGCGTTGGCTACCGGTGATATGCTGAAGAGCTTTGGCGGGAATGGGCTCAACCGCTT	1105		
DB	691	CATAGCGTTGGCTACCGGTGATATGCTGAAGAGCTTTGGCGGGAATGGGCTCAACCGCTT	750		
QY	1106	CCTCGTGCTTTACGGTATCCCGCTCCGATTCGACGCGCATCGCTTCTATCGCCTTCT	1165		
DB	751	CCTCGTGCTTTACGGTATCCCGCTCCGATTCGACGCGCATCGCTTCTATCGCCTTCT	810		
QY	1166	TGACGAGTTCTTCTGAG	1182		
DB	811	TGACGAGTTCTTCTGAG	827		

RESULT 2	804 bp	linear	EST 02-AUG-2004
CK291519	EST754233	Nicotiana benthamiana mixed tissue	CDNA library
LOCUS	normalized, full-length Nicotiana benthamiana cDNA clone NEMC276 5'		
DEFINITION	end, mRNA sequence.		
ACCESSION	CK291519		
VERSION	CK291519.1	GI:39872047	

KEYWORDS	EST.
SOURCE	Nicotiana benthamiana
ORGANISM	Nicotiana benthamiana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliopsida; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Nicotiana. 1 (bases 1 to 804)
REFERENCE	Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A., Day,B., Staskawicz,B., Jin,H. and Baker,B. Generation of EST sequences from the University of Arizona Genomics Institute via http://genome.arizona.edu/orders/ . Seq primer: ATT TAG GTG ACA CTA TGA TAG.
TITLE	Location/Qualifiers
JOURNAL	1. 804
COMMENT	/organism="Nicotiana benthamiana" /mol_type="mRNA" /db_xref="taxon:4100" /clone="NBMC276" /tissue type="abiotic and biotic stress-treated leaves, callus tissue and root tissue" /lab_host="DH10B-Tona" /clone_lib="Nicotiana benthamiana mixed tissue cdna library, normalized, full-length" /notes=Vector: pCMVSPORT6.1; Site:1: EcoRI; Site_2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."
FEATURES	
source	
ORIGIN	
Query Match	9.9%; Score 794; DB 7; Length 804;
Best Local Similarity	100.0%; Pred. No. 7.5e-174;
Matches	794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	389 GATTGACAGATGGATTCCAGCAGAGTTCTCCGCCCGCTTGGTGAGAGGCTATTTCG 448
Db	
2	GATTGACAGATGGATTGCACAGAGTTCTCCGCCCGCTTGGTGAGAGGCTATTTCG 61
QY	449 CTATGACTGGGCAACAAGACAAATCGGCTGTCTGATCGCGCGCTTTCGGCTGTCA 508
Db	
62	CTATGACTGGGCAACAAGACAAATCGGCTGTCTGATCGCGCGCTTTCGGCTGTCA 121
QY	509 GCAGGGGGCCCGGTTCTTTTGTCAAGACCGACCTTCGGGTGCCCTGAATCAATGCA 568
Db	
122	GCAGGGGGCCCGGTTCTTTTGTCAAGACCGACCTTCGGGTGCCCTGAATCAATGCA 191
QY	569 GGACGAGGGACGCCGGCTATCGTGGCTGGCCACGACGGCGCTTCCTTTGGCGAGCTGTG 628
Db	
182	GGACGAGGGACGCCGGCTATCGTGGCTGGCCACGACGGCGCTTCCTTTGGCGAGCTGTG 241
QY	629 CGACGTTGTCACTGAAGCGGGAAGGACGTGGCTGTATTGGGGCGAAGTCCGGGGCAG 688
Db	
242	CGACGTTGTCACTGAAGCGGGAAGGACGTGGCTGTATTGGGGCGAAGTCCGGGGCAG 301
QY	689 TCCTCTGTCACTCATCTTCTCTCCGCGAGAAGTAGTCCATCATGGCTGATGCAATGCG 748
Db	
302	TCCTCTGTCACTCATCTTCTCTCCGCGAGAAGTAGTCCATCATGGCTGATGCAATGCG 361
QY	749 GCGGCTGCATACGCTTTCATCCGGCTACCTGCCCATTCGACCACAAGCGAAAACATTCG 808
Db	
362	GCGGCTGCATACGCTTTCATCCGGCTACCTGCCCATTCGACCACAAGCGAAAACATTCG 421

QY 809 CGAGGACGACGTAATCGGATGGAGCCGGTCTTGTGATCAGGATGATCTCGACGAGA 868  
 |||||  
 Db 422 CGAGGACGACGTAATCGGATGGAGCCGGTCTTGTGATCAGGATGATCTGGACGAGA 481  
 |||||  
 QY 869 GCATCAGGGGTCGCGCCAGCCGCAACTGTTGCCAGGCTCAAGGCGCGCATGCCGACGG 928  
 |||||  
 Db 482 GCATCAGGGGTCGCGCCAGCCGCAACTGTTGCCAGGCTCAAGGCGCGCATGCCGACGG 541  
 |||||  
 QY 929 CGAGGATCTCTGTGACCCATGGCGATGCTGCTTGGCGCAATATCATGGTGGAAAAATGG 988  
 |||||  
 Db 542 CGAGGATCTCTGTGACCCATGGCGATGCTGCTTGGCGCAATATCATGGTGGAAAAATGG 601  
 |||||  
 QY 989 CCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGGCGGACCCCTATCAGGACAT 1048  
 |||||  
 Db 602 CCGCTTTTCTGATTCATCGACTGTGGCGGCTGGGTGGCGGACCCCTATCAGGACAT 661  
 |||||  
 QY 1049 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCC 1108  
 |||||  
 Db 662 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCC 721  
 |||||  
 QY 1109 GGTGCTTTACGGTATCGCGCTCCGATTCGAGCGGATCGCCTTCTATCGCCTTTTGA 1168  
 |||||  
 Db 722 GGTGCTTTACGGTATCGCGCTCCGATTCGAGCGGATCGCCTTCTATCGCCTTTTGA 781  
 |||||  
 QY 1169 CGAGTCTCTCTGAG 1182  
 |||||  
 Db 782 CGAGTCTCTCTGAG 795  
 |||||

## RESULT 3

CK287297

## LOCUS

DEFINITION CK287297 856 bp mRNA linear EST 02-AUG-2004  
 EST750019 Nicotiana benthamiana mixed tissue cDNA library,  
 normalized, full-length Nicotiana benthamiana cDNA clone NEMB815 5'  
 end, mRNA sequence.

## ACCESSION

CK287297

## VERSION

CK287297.1

## KEYWORDS

GI:39863696

## SOURCE

EST.

## ORGANISM

Nicotiana benthamiana

## REFERENCE

AUTHORS

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
 Staskiewicz, B., Jin, H. and Baker, B.

Generation of EST sequences from Nicotiana benthamiana

Unpublished (2003)

Other ESTs: EST750020

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics

Institute via <http://genome.arizona.edu/orders/>

Seq primer: ATT TAG GTG ACA CTA TAG.

## FEATURES

source

1..856

Location/Qualifiers

/organism="Nicotiana benthamiana"

/mol\_type="mRNA"

/db\_xref="taxon:4100"

/clone="NEMB815"

/tissue\_type="abiotic and biotic stress-treated leaves,"

callus tissue and root tissue"

/lab\_host="DH10B-Tona"

/clone\_lib="Nicotiana benthamiana mixed tissue cDNA

library, normalized, full-length"

/note="vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Nicotiana benthamiana

tissues that include callus, roots from liquid culture

grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),

cold-stressed leaves (5 C 3 hr, 6hr), and pathogen

challenged leaves (*Pseudomonas syringae* pv *tomato* 12 hr;  
*Xanthomonas campestris* pv *campestris* 12 hr, 18hr;  
*Pseudomonas syringae* pv *phaseolicola* 18hr, and *Xanthomonas*  
*campestris* pv *vesicatoria* 18hr). RNA was isolated from  
 these tissues and pooled in approximately equal molar  
 amounts."

## ORIGIN

Query Match 9.9%; Score 794; DB 7; Length 856;  
 Best Local Similarity 100.0%; Pred. No. 7.6e-174;  
 Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 389 GATTGAACAAGATGGATGTCAGCAGGTTCTCCGGCCGCTTGGGTGGAGAGCTATTTCGG 448  
 |||||  
 Db 14 GATTGAACAAGATGGATGTCAGCAGGTTCTCCGGCCGCTTGGGTGGAGAGCTATTTCGG 73  
 |||||  
 QY 449 CTATGACTGGGCACAAACAGACAATCGGCTGCTCTGATGCCCGCGTGTTCGGCTGTACG 508  
 |||||  
 Db 74 CTATGACTGGGCACAAACAGACAATCGGCTGCTCTGATGCCCGCGTGTTCGGCTGTACG 133  
 |||||  
 QY 509 GCAGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTTGAATGAATGCA 568  
 |||||  
 Db 134 GCAGGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCCGGTGCCTTGAATGAATGCA 193  
 |||||  
 QY 569 GGACGAGGACGCGCGCTATCGTGGCTGGCCACAGCGGCGTTCCTTGGCAGCTGTGCT 628  
 |||||  
 Db 194 GGACGAGGACGCGCGCTATCGTGGCTGGCCACAGCGGCGTTCCTTGGCAGCTGTGCT 253  
 |||||  
 QY 629 CGAGCTTGTCTCACTGAAGCGGGAAGGACTGCTCTCTATTTGGGCGAAGTCCGCTGATGCAATCG 748  
 |||||  
 Db 314 TCCTCTGTCTCTCACTGACCTTCTCTCTCGGAGAAAGTATCCATCATGCTGATGCAATCG 373  
 |||||  
 QY 749 CGGCTCGATACGCTTGTATCCGGCTACTGCTCCCATTCGACCAACCAAGCAAAATCGCAT 808  
 |||||  
 Db 374 CGGCTCGATACGCTTGTATCCGGCTACTGCTCCCATTCGACCAACCAAGCAAAATCGCAT 433  
 |||||  
 QY 809 CGAGGACGACGTAATCGGATGGAGCGGCTTGTGTCGATCAGGATGATCTGGACGAGA 868  
 |||||  
 Db 434 CGAGGACGACGTAATCGGATGGAGCGGCTTGTGTCGATCAGGATGATCTGGACGAGA 493  
 |||||  
 QY 869 GCATCAGGGGTCGCGCCAGCCGAACTGTTCCGCGAGCTCAAGCGCGCATGCCCGACGG 928  
 |||||  
 Db 494 GCATCAGGGGTCGCGCCAGCCGAACTGTTCCGCGAGCTCAAGCGCGCATGCCCGACGG 553  
 |||||  
 QY 929 CGAGGATCTCTGTGTCGACCCATGCGCATGCTGCTTCCGGAATATCATGGTGGAAAAATGG 988  
 |||||  
 Db 554 CGAGGATCTCTGTGTCGACCCATGCGCATGCTGCTTCCGGAATATCATGGTGGAAAAATGG 613  
 |||||  
 QY 989 CCGCTTTTCTGGAATTCATCGACTGTGCCCGGCTGGGTGGCGACCGCTATCAGGACAT 1048  
 |||||  
 Db 614 CCGCTTTTCTGGAATTCATCGACTGTGCCCGGCTGGGTGGCGACCGCTATCAGGACAT 673  
 |||||  
 QY 1049 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCC 1108  
 |||||  
 Db 674 AGCGTTGGCTACCGGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCC 733  
 |||||  
 QY 1109 CGTGTCTTACGGTATCGCCGCTCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGA 1168  
 |||||  
 Db 734 CGTGTCTTACGGTATCGCCGCTCCGATTCGAGCGCATCGCTTCTATCGCTTCTTGA 793  
 |||||  
 QY 1169 CGAGTCTCTCTGAG 1182  
 |||||  
 Db 794 CGAGTCTCTCTGAG 807  
 |||||

## RESULT 4

CK287930

## LOCUS

DEFINITION

CK287930 910 bp mRNA linear EST 02-AUG-2004

EST750652 Nicotiana benthamiana mixed tissue cDNA library,

normalized, full-length Nicotiana benthamiana cDNA clone NBMC75 5'  
end, mRNA sequence.  
CK287930.1 GI:39864940  
EST.  
SOURCE  
ORGANISM  
Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 910)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
Staskawicz, B., Jin, H. and Baker, B.  
Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)  
Other ESTs: EST750653  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1..910  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBMC75"  
/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Nicotiana benthamiana  
tissues that include callus, roots from liquid culture  
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
challenged leaves (Pseudomonas syringae pv tomato 12 hr;  
Xanthomonas campestris pv campestris 12 hr, 18hr;  
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas  
campestris pv vesicatoria 18hr). RNA was isolated from  
these tissues and pooled in approximately equal molar  
amounts."  
ORIGIN  
Query Match 9.9%; Score 794; DB 7; Length 910;  
Best Local Similarity 100.0%; Pred. No. 7.7e-174;  
Matches 794; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 389 GATTGAACAGATGATGACGAGGTTCTCGGCGGCTGGGTGAGAGGCTATTCGG 448  
Db 4 GATTGAACAGATGATGACGAGGTTCTCGGCGGCTGGGTGAGAGGCTATTCGG 63  
Qy 449 CTATGACTGGGCAACAGACAAATCGGCTGCTGTGATGCCCGCTTTCGGCTTCAGC 508  
Db 64 CTATGACTGGGCAACAGACAAATCGGCTGCTGTGATGCCCGCTTTCGGCTTCAGC 123  
Qy 509 GCAGGGGCGCGCGGTTCTTTTGTCAAGACCGACTGTCCGGTGCCTGAATGAATGCA 568  
Db 124 GCAGGGGCGCGGTTCTTTTGTCAAGACCGACTGTCCGGTGCCTGAATGAATGCA 183  
Qy 569 GGACGAGGACGCGGCTATCGTGGTGGGCAACGCGGCTTCTTGGCGAGCTGTGCT 628  
Db 184 GGACGAGGACGCGGCTATCGTGGTGGGCAACGCGGCTTCTTGGCGAGCTGTGCT 243  
Qy 629 CGAGCTTGTCACTGAAGCGGAGGAGTGTGCTCTATTGGGCGAAGTGGCGGGGACGGA 688  
Db 244 CGAGCTTGTCACTGAAGCGGAGGAGTGTGCTCTATTGGGCGAAGTGGCGGGGACGGA 303  
Qy 689 TCTCTCTCATCTCACCTTGTCTCTGCGGAGAAATATCCATCATGCTGATGCAATGCG 748

Db 304 TCTCTCTCATCTCACCTTGTCTCTGCGGAGAAATATCCATCATGCTGATGCAATGCG 363  
Qy 749 GCGGCTGCATACGCTTGTATCGCGGTACCTGCCATTTCGACCAACGAGAAACATCGCAT 808  
Db 364 GCGGCTGCATACGCTTGTATCGCGGTACCTGCCATTTCGACCAACGAGAAACATCGCAT 423  
Qy 809 CGAGCGAGCACGTACTTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGA 868  
Db 424 CGAGCGAGCACGTACTTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTGGACGAAGA 483  
Qy 869 GCATCAGGGGCTCGCGCAGCCGAACTGTTCCGAGGCTCAAGCGCGCATGCCGACCG 928  
Db 484 GCATCAGGGGCTCGCGCAGCCGAACTGTTCCGAGGCTCAAGCGCGCATGCCGACCG 543  
Qy 929 CGAGGATCTCGTGTGACCCATGGCGATGCTGCTTCCGAAATATCATGCTGGAATATGG 988  
Db 544 CGAGGATCTCGTGTGACCCATGGCGATGCTGCTTCCGAAATATCATGCTGGAATATGG 603  
Qy 989 CCGCTTTTCTGGATTCATCGATGCGCGGTGGTGGCGGACCGCTATCAGGACAT 1048  
Db 604 CCGCTTTTCTGGATTCATCGATGCGCGGTGGTGGCGGACCGCTATCAGGACAT 663  
Qy 1049 AGCGTTGGCTACCGCTGATATTCCTCAGAGGCTTGGCGGAGATGGCTGACCGCTTCT 1108  
Db 664 AGCGTTGGCTACCGCTGATATTCCTCAGAGGCTTGGCGGAGATGGCTGACCGCTTCT 723  
Qy 1109 CGTGTCTTTACGGTATCGCGCTCCCGATTTCGAGCGCATCGCTTCTATCGCTTCTTGA 1168  
Db 724 CGTGTCTTTACGGTATCGCGCTCCCGATTTCGAGCGCATCGCTTCTATCGCTTCTTGA 783  
Qy 1169 CGAGTTCTTCTGAG 1182  
Db 784 CGAGTTCTTCTGAG 797  
CK291799 933 bp mRNA linear EST 02-AUG-2004  
EST754513 Nicotiana benthamiana mixed tissue cDNA library,  
normalized, full-length Nicotiana benthamiana cDNA clone NBMC477 5'  
end, mRNA sequence.  
CK291799  
CK291799.1 GI:39872608  
EST.  
SOURCE  
ORGANISM  
Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
Staskawicz, B., Jin, H. and Baker, B.  
Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)  
Other ESTs: EST754514  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via http://genome.arizona.edu/orders/  
Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1..933  
/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBMC477"  
/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"  
FEATURES  
source

/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: RNA was isolated from Nicotiana benthamiana tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

ORIGIN	Query Match		Best Local Similarity		Matches		Score		DB 7: Length		933;		Pred. No. 7.7e-174;		Mismatches		0;		Indels		0;		Gaps	
	Conservative		0;		0;		0;		0;		0;		0;		0;		0;		0;		0;		0;	
QY	389	GATTGAACAAGATGCGGCTTCTTCTGCGGCGCTTGGTGAGAGGCTATTCGG	448																					
Db	46	GATTGAACAAGATGCGGCTTCTTCCGCGCGCTTGGTGAGAGGCTATTCGG	105																					
QY	449	CTATGACTGGGACACAGCAATCGGCTGCTCTGATCGCGCGGCTTCCGGCTGTCAGC	508																					
Db	106	CTATGACTGGGACACAGCAATCGGCTGCTCTGATCGCGCGGCTTCCGGCTGTCAGC	165																					
QY	509	GCAGGGGCGCGCGCTTCTTCTGCAAGACCGACCTGTCGGTGCCTGGAATGACTGCA	568																					
Db	166	GCAGGGGCGCGCGCTTCTTCTGCAAGACCGACCTGTCGGTGCCTGGAATGACTGCA	225																					
QY	569	GGAAGGACGCGGCTTCTGCTGCGCACGACGCGCGCTTCTTCCGCGAGCTGTGCT	628																					
Db	226	GGACGAGGACGCGGCTTCTGCTGCGCACGACGCGCGCTTCTTCCGCGAGCTGTGCT	285																					
QY	629	CGAGTTGTCACTCAAGCGGGAAGGACTGCTGCTATGCGGCAAGTCCGGGCGAGGA	688																					
Db	286	CGAGTTGTCACTCAAGCGGGAAGGACTGCTGCTATGCGGCAAGTCCGGGCGAGGA	345																					
QY	689	TCTCTGTCTATCTACCTTGTCTGCGGAGAAAGTATCCATCATGGCTGATGAATGCG	748																					
Db	346	TCTCTGTCTATCTACCTTGTCTGCGGAGAAAGTATCCATCATGGCTGATGAATGCG	405																					
QY	749	GCGGCTGCATACGCTTGTATCGGCTTACCTGCGCCATTCGACCACCAAGCAACATCGCAT	808																					
Db	406	GCGGCTGCATACGCTTGTATCGGCTTACCTGCGCCATTCGACCACCAAGCAACATCGCAT	465																					
QY	809	CGAGCGAGCAGTACTCGGATGGAAGCGGCTTGTGATCAGGATGATCTGGAAGA	868																					
Db	466	CGAGCGAGCAGTACTCGGATGGAAGCGGCTTGTGATCAGGATGATCTGGAAGA	525																					
QY	869	GCATCAGGGGCTCGCGCCAGCGCAACTGTTCGCCAGGCTCAAGCGCGCATGCCGACGG	928																					
Db	526	GCATCAGGGGCTCGCGCCAGCGCAACTGTTCGCCAGGCTCAAGCGCGCATGCCGACGG	585																					
QY	929	CGAGGATCTCGTCTGTGACCCATGCGATGCTGCTTCCGCAATATCATGTGGAAATGG	988																					
Db	586	CGAGGATCTCGTCTGTGACCCATGCGATGCTGCTTCCGCAATATCATGTGGAAATGG	645																					
QY	989	CCGCTTTTCTGATTCATGCACTGTGCGCGCTGGTGTGCGCAACCGCTATCAGGACAT	1048																					
Db	646	CCGCTTTTCTGATTCATGCACTGTGCGCGCTGGTGTGCGCAACCGCTATCAGGACAT	705																					
QY	1049	AGCGTTGGCTACCGCTGATATTTGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCT	1108																					
Db	706	AGCGTTGGCTACCGCTGATATTTGCTGAAGAGCTTGGCGGCAATGGGCTGACCGCTTCT	765																					
QY	1109	CGTGCTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGA	1168																					
Db	766	CGTGCTTTACGGTATCGCGCTCCCGATTCGACGCGCATCGCCTTCTATCGCCTTCTTGA	825																					
QY	1169	CGAGTTCTTCTGAG	1182																					
Db	826	CGAGTTCTTCTGAG	839																					

# RESULT 6

CK256977

LOCUS

DEFINITION

CK256977

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

location/Qualifiers

1. .936

/organism="Solanum tuberosum"

/mol\_type="mRNA"

/cultivar="Kennebec"

/db\_xref="taxon:4113"

/clone="POCD170"

/tissue\_type="callus"

/lab\_host="DH10B-Tona"

/clone\_lib="potato callus cDNA library, normalized and

/note="length"

/supplier: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;

supplier: RNA was isolated from Solanum tuberosum var.

Kennebec callus tissue grown on solid media."

ORIGIN

Query Match

Best Local Similarity

Matches

794; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

QY

389

GATTGAACAAGATGCGGCTTCTTCCGCGCGCTTGGTGAGAGGCTATTCGG

448

Db

14

GATTGAACAAGATGCGGCTTCTTCCGCGCGCTTGGTGAGAGGCTATTCGG

73

QY

449

CTATGACTGGGCAACAGCAATCGGCTGCTGTGATGCGCGCGCTTCCGGCTGTGAGC

508

Db

74

CTATGACTGGGCAACAGCAATCGGCTGCTGTGATGCGCGCGCTTCCGGCTGTGAGC

133

QY

509

GCAGGGGCGCGCGCTTCTTCCGCGCGCTTGGTGAGAGGCTATTCGGCTGTGAGC

568

Db

134

GCAGGGGCGCGCGCTTCTTCCGCGCGCTTGGTGAGAGGCTATTCGGCTGTGAGC

193

QY

569

GGACGAGGCGCGCGCTTCTTCCGCGCGCTTGGTGAGAGGCTATTCGGCTGTGAGC

628

Db

194

GGACGAGGCGCGCGCTTCTTCCGCGCGCTTGGTGAGAGGCTATTCGGCTGTGAGC

253

QY

629

CGACGTTGTCTCACTGAACGCGGAAAGGACTGGCTGCTATTGGCGGAAAGTGGCGGCGAGGA

688

Db

254

CGACGTTGTCTCACTGAACGCGGAAAGGACTGGCTGCTATTGGCGGAAAGTGGCGGCGAGGA

313

QY

689

TCTCCCTGTCTCATCTCAGCTTCTCTCCGAGAAAGTATCCATCATGCTGATGCAATGCG

748

Db

314

TCTCCCTGTCTCATCTCAGCTTCTCTCCGAGAAAGTATCCATCATGCTGATGCAATGCG

373

QY

749

GCAGGCTGCATACGCTTGTATCCGCGCTACCTGCCCATTCGACACCAAGCGAAACATCGCAT

808

Db

374

GCAGGCTGCATACGCTTGTATCCGCGCTACCTGCCCATTCGACACCAAGCGAAACATCGCAT

433







tissues that include callus, roots from liquid culture grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr), cold-stressed leaves (5 C 3 hr, 6hr), and pathogen challenged leaves (Pseudomonas syringae pv tomato 12 hr; Xanthomonas campestris pv campestris 12 hr, 18hr; Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas campestris pv vesicatoria 18hr). RNA was isolated from these tissues and pooled in approximately equal molar amounts."

## ORIGIN

Query Match 9.7%; Score 778.8; DB 7; Length 811;  
Best Local Similarity 99.7%; Pred. No. 2.6e-170; Indels 0; Gaps 0;  
Matches 780; Conservative 0; Mismatches 2;

QY 386 CATGATTGAACAGATGGATTGCACGAGTTCTCCGGCGCTGGGTGAGAGGTATT 445  
DB 30 CATGATTGAACAGATGGATTGCACGAGTTCTCCGGCGCTGGGTGAGAGGTATT 89

QY 446 CGGCTATGACTGGGACACAGACAATCGCTGCTGATGCCGCGTGTCCGGCTGTC 505  
DB 90 CGGCTATGACTGGGACACAGACAATCGCTGCTGATGCCGCGTGTCCGGCTGTC 149

QY 506 AGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCGGTGCCCTGAATGAAC 565  
DB 150 AGCGCAGGGCGCCCGGTTCTTTTGTCAAGACCGACCTGTCGGTGCCCTGAATGAAC 209

QY 566 GCAGGACGAGGACGCGGCTATCGTGGCTGGCCACGACGGCGTTCCTTGGCGAGCTGT 625  
DB 210 GCAGGACGAGGACGCGGCTATCGTGGCTGGCCACGACGGCGTTCCTTGGCGAGCTGT 269

QY 626 GCTCGAGCTTGTCACTGAAGCGGAGGAGTGGCTGCTATTGGCGGAGTGGCGGGCA 685  
DB 270 GCTCGAGCTTGTCACTGAAGCGGAGGAGTGGCTGCTATTGGCGGAGTGGCGGGCA 329

QY 686 GGATCTCTGCTATCTCACCTTGCTCTCGCGAGAAAGTATCCATCATGCTGATGCAAT 745  
DB 330 GGATCTCTGCTATCTCACCTTGCTCTCGCGAGAAAGTATCCATCATGCTGATGCAAT 389

QY 746 GCGGCGGCTGCATACGCTTGCATCGGCTACCTGCCCATTCGACCAACGCAAAATCG 805  
DB 390 GCGGCGGCTGCATACGCTTGCATCGGCTACCTGCCCATTCGACCAACGCAAAATCG 449

QY 806 CATCGAGCGAGCAGTACTCGATGGAGCGGCTTGTGATCAGATGATCTGGACGA 865  
DB 450 CATCGAGCGAGCAGTACTCGATGGAGCGGCTTGTGATCAGATGATCTGGACGA 509

QY 866 AGAGCATCAGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGCGCGCATGCCGA 925  
DB 510 AGAGCATCAGGGCTCGCGCAGCGCAACTGTTGCGCAGGCTCAAGCGCGCATGCCGA 569

QY 926 CGGCGAGGATCTCGTGTGACCCATGCGGATGCTGTTGCCGAATATCATGTGGAAAA 985  
DB 570 CGGCGAGTATCTCGTGTGACCCATGCGGATGCTGTTGCCGAATATCATGTGGAAAA 629

QY 986 TGGCGGCTTTCTGGATTTCATCGACTGTCGGCGGCTGGGTGTGCGGACCGCTATCAGGA 1045  
DB 630 TGGCGGCTTTCTGGATTTCATCGACTGTCGGCGGCTGGGTGTGCGGACCGCTATCAGGA 689

QY 1046 CATAGCCTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGGGCAATGGGCTGACCGCTT 1105  
DB 690 CATAGCCTTGGCTACCGCTGATATTGCTGAAGAGCTTGGCGGGCAATGGGCTGACCGCTT 749

QY 1106 CCTCGTCTTACGGTATGCGCGCTCCCGATTCGACGCGATCGCCCTTCTATGCGCTTCT 1165  
DB 750 CCTCGTCTTACGGTATGCGCGCTCCCGATTCGACGCGATCGCCCTTCTATGCGCTTCT 809

QY 1166 TG 1167  
DB 810 TG 811

RESULT 10

CK288711  
LOCUS  
DEFINITION

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

FEATURES  
source

CK288711 878 bp mRNA linear EST 02-AUG-2004  
EST751433 Nicotiana benthamiana mixed tissue cDNA library.  
normalized, full-length Nicotiana benthamiana cDNA clone NBMBI49 5'  
end, mRNA sequence.  
CK288711  
CK288711.1 GI:39866496  
EST.  
Nicotiana benthamiana  
Nicotiana benthamiana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Nicotiana.  
1 (bases 1 to 878)  
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A., Day, B.,  
Taskawicz, B., Jin, H., and Baker, B.  
Generation of EST sequences from Nicotiana benthamiana  
Unpublished (2003)  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from the University of Arizona Genomics  
Institute via <http://genome.arizona.edu/orders/>.  
Seq primer: ATT TAG GTG ACA CTA TAG.

Location/Qualifiers  
1..878

/organism="Nicotiana benthamiana"  
/mol\_type="mRNA"  
/db\_xref="taxon:4100"  
/clone="NBMBI49"  
/tissue\_type="abiotic and biotic stress-treated leaves,  
callus tissue and root tissue"  
/lab\_host="DH10B-Tona"  
/clone\_lib="Nicotiana benthamiana mixed tissue cDNA  
library, normalized, full-length"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Nicotiana benthamiana  
tissues that include callus, roots from liquid culture  
grown plants, heat-stressed leaves (38 C, 3 hr and 6 hr),  
cold-stressed leaves (5 C 3 hr, 6hr), and pathogen  
challenged leaves (Pseudomonas syringae pv tomato 12 hr;  
Xanthomonas campestris pv campestris 12 hr, 18hr;  
Pseudomonas syringae pv phaseolicola 18hr, and Xanthomonas  
campestris pv vesicatoria 18hr). RNA was isolated from  
these tissues and pooled in approximately equal molar  
amounts."

## ORIGIN

Query Match 7.3%; Score 587; DB 7; Length 878;  
Best Local Similarity 100.0%; Pred. No. 1.5e-125; Indels 0; Gaps 0;  
Matches 587; Conservative 0; Mismatches 0;

QY 596 GGCACGACGCGGCTTCCTTGCAGCTGTGCTCGAGTGTGTCACTGAAGCGGAAGGGA 655  
DB 1 GGCACGACGCGGCTTCCTTGCAGCTGTGCTCGAGTGTGTCACTGAAGCGGAAGGGA 60

QY 656 CTGGCTGCTATTGGCGAAGTGGCGGAGGATCTCTGTCATCTCACCTTGTCTCTGC 715  
DB 61 CTGGCTGCTATTGGCGAAGTGGCGGAGGATCTCTGTCATCTCACCTTGTCTCTGC 120

QY 716 CGAAGAAAGTATCATCATGCTGATGCAATGCGCGGCTGCATACGCTTGTATCCGGCTAC 775  
DB 121 CGAAGAAAGTATCATCATGCTGATGCAATGCGCGGCTGCATACGCTTGTATCCGGCTAC 180

QY 776 CTGCCATTCCGACCAACGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGGAAGC 835  
DB 181 CTGCCATTCCGACCAACGCGAAACATCGCATCGAGCGAGCAGTACTCGGATGGGAAGC 240

QY 836 CGGTCTTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGTCCGCCAGCCGAACT 895  
DB 241 CGGTCTTGTGATCAGGATGATCTGGACGAGAGCATCAGGGGTCCGCCAGCCGAACT 300

```

QY      896 GTTCGCCAGGCTCAAGCGCGCATGCCGACGGGAGGATCTCGTGTGACCCATGGCGA 955
      |||
Db      301 GTTCGCCAGGCTCAAGCGCGCATGCCGACGGGAGGATCTCGTGTGACCCATGGCGA 360
      |||
QY      956 TGCGTGTCTGCCGAATATCATGTGTGGAATAATGGCCGCTTTTCTGGAATCATCGACTGTGG 1015
      |||
Db      361 TGCGTGTCTGCCGAATATCATGTGTGGAATAATGGCCGCTTTTCTGGAATCATCGACTGTGG 420
      |||
QY      1016 CCGGCTGGGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCGTGATATGTCTGA 1075
      |||
Db      421 CCGGCTGGGTGGCGGACCGCTATCAGGACATAGCTTGGCTACCGTGATATGTCTGA 480
      |||
QY      1076 AGAGCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTTACCGTATCGCGCTCCCGA 1135
      |||
Db      481 AGAGCTTGGCGGGAATGGGCTGACCGCTTCTCGTCTTTTACCGTATCGCGCTCCCGA 540
      |||
QY      1136 TTGCGAGCGCATCGCCTTCTATCGCTTCTTGACGAGTCTTCTGAG 1182
      |||
Db      541 TTGCGAGCGCATCGCCTTCTATCGCTTCTTGACGAGTCTTCTGAG 587
      |||

RESULT 11
AQ361914
LOCUS
DEFINITION
  mgxb005K01f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
  clone mgxb005K01f, genomic survey sequence.
ACCESSION
  AQ361914
VERSION
  AQ361914.1 GI:4211753
KEYWORDS
  GSS.
ORGANISM
  Magnaporthe grisea (anamorph: Pyricularia grisea)
  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
  Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
  1 (bases 1 to 789)
AUTHORS
  Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
  Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
  A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
  Genome
  Unpublished (1998)
  Contact: Dean RA
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson University, Clemson, SC 29634
  Tel: 864 656 5737
  Fax: 864 656 4293
  Email: rdean@clemson.edu
  Seq primer: TAATACGACTCCTATAGGG
  Class: BAC ends
  High quality sequence start: 41
  High quality sequence stop: 392.
  Location/Qualifiers
    1..789
      /organism="Magnaporthe grisea"
      /mol_type="genomic DNA"
      /strain="70-15"
      /db_xref="taxon:148305"
      /clone="mgxb005K01f"
      /tissue_type="Protoplasts"
      /lab_host="E. coli DH10B"
      /notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
      Rice blast is one of the most devastating fungal diseases
      of rice world wide. It is a filamentous ascomycete with
      a haploid genome (n=7) of approximately 40 Mbp. Rice
      blast is an important model fungal pathogen for studying
      numerous aspects of the fungal-host interaction. In
      order to facilitate genome wide analysis, a BAC library
      containing 9216 clones with an average insert size of 130
      kbp was constructed. This library represents greater
      than 25x genome coverage. High density colony filters
      are available upon request."
  
```

ORIGIN

```

Query Match      7.2%; Score 577.4; DB 8; Length 789;
Best Local Similarity 97.8%; Pred. No. 2.6e-123;
Matches 584; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

QY      581 GCGCCTATCGTGGCTGCGCACGACGGGCGTTCTTTCGCGACGCTGCTGCACGTTGTTCAC 640
      |||
Db      39  GCGCTATCTGCTGCTGCGCACGACGGGCGTTCTTTCGCGACGCTGCTGCACGTTGTTCAC 98
      |||
QY      641 TGAAGCGGGGAAGGACTGTGGCTGTATTGGGCGAAGTGCOCGGGGCAGGATCTCTGTCTATC 700
      |||
Db      99  TGAAGCGGGGAAGGACTGTGGCTGTATTGGGCGAAGTGCOCGGGGCAGGATCTCTGTCTATC 158
      |||
QY      701 TCACCTTGTCTCTCGCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGTCATAC 760
      |||
Db      159 TCACCTTGTCTCTCGCGAGAAAGTATCCATCATGGCTGATGCAATGCGGCGGCTGTCATAC 218
      |||
QY      761 GCTTGATCCCGGCTACCTGCCCATTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCAGC 820
      |||
Db      219 GCTTGATCCCGGCTACCTGCCCATTCGACCAACCAAGCGAAACATCGCATCGAGCGAGCAGC 278
      |||
QY      821 TACTCGGATGGAAGCGCGTCTTGTGATCAGGATGATCTGACGAGAGATCAGGGGCT 880
      |||
Db      279 TACTCGGATGGAAGCGCGTCTTGTGATCAGGATGATCTGACGAGAGATCAGGGGCT 338
      |||
QY      881 CCGCGCAGCGAAGTGTTCGCGAGGCTCAAGGCGCGCATGCCGACGGCGAGGATCTCGT 940
      |||
Db      339 CCGCGCAGCGAAGTGTTCGCGAGGCTCAAGGCGCGCATGCCGACGGCGAGGATCTCGT 398
      |||
QY      941 CGTGACCCATGGCGATGCCCTGTTCGGAATATCATGTTGGAATAATGGCGCTTTCTTCTGG 1000
      |||
Db      399 CGTGACCCATGGCGATGCCCTGTTCGGAATATCATGTTGGAATAATGGCGCTTTCTTCTGG 458
      |||
QY      1001 ATTTCATGCACTGTGGCGCGCTGGGTGTGGCGACCGCTATCAGGACATACGTTGGCTAC 1060
      |||
Db      459 ATTTCATGCACTGTGGCGCGCTGGGTGTGGCGACCGCTATCAGGACATACGTTGGCTAC 518
      |||
QY      1061 CCGTGATATGCTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCTGCTTACGG 1120
      |||
Db      519 CCGTGATATGCTGAAGAGCTTGGCGCGAATGGGCTGACCGCTTCTCTGCTTACGG 578
      |||
QY      1121 TATCGCGCTCCCGATTTCGCGAGCGCATCGCTTCTTATCGCGCTTCTTACGAGTTCTT 1177
      |||
Db      579 TATCGCGCTCCCGATTTCGCGAGCGCATCGCTTCTTATCGCGCTTCTTACGAGTTCTT 635
      |||

RESULT 12
AQ398387
LOCUS
DEFINITION
  mgxb000J05f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
  clone mgxb000J05f, genomic survey sequence.
ACCESSION
  AQ398387
VERSION
  AQ398387.1 GI:4369414
KEYWORDS
  GSS.
SOURCE
  Magnaporthe grisea (anamorph: Pyricularia grisea)
  Magnaporthe grisea
  Unpublished (1998)
  Contact: Dean RA
  Clemson University Genomics Institute
  Clemson University
  100 Jordan Hall, Clemson University, Clemson, SC 29634
  Tel: 864 656 5737
  Fax: 864 656 4293
  Email: rdean@clemson.edu
  Seq primer: TAATACGACTCCTATAGGG
  Class: BAC ends
  High quality sequence start: 41
  High quality sequence stop: 392.
  Location/Qualifiers
    1..789
      /organism="Magnaporthe grisea"
      /mol_type="genomic DNA"
      /strain="70-15"
      /db_xref="taxon:148305"
      /clone="mgxb000J05f"
      /tissue_type="Protoplasts"
      /lab_host="E. coli DH10B"
      /notes="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
      Rice blast is one of the most devastating fungal diseases
      of rice world wide. It is a filamentous ascomycete with
      a haploid genome (n=7) of approximately 40 Mbp. Rice
      blast is an important model fungal pathogen for studying
      numerous aspects of the fungal-host interaction. In
      order to facilitate genome wide analysis, a BAC library
      containing 9216 clones with an average insert size of 130
      kbp was constructed. This library represents greater
      than 25x genome coverage. High density colony filters
      are available upon request."
  
```

```
FEATURES
  source      Location/Qualifiers
    1..620
      /organism="Magnaporthe grisea"
      /mol_type="genomic DNA"
      /strain="70-15"
      /db_xref="taxon:148305"
      /clone="mgxb0009J05f"
      /tissue_type="Protoplasts"
      /lab_host="E. coli DH10B"
      /clone_lib="CUGI Rice Blast BAC Library"
      /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
Rice blast is one of the most devastating fungal diseases
of rice world wide. It is a filamentous ascomycete with
a haploid genome (n=7) of approximately 40 Mbp. Rice
blast is an important model fungal pathogen for studying
numerous aspects of the fungal-host interaction. In
order to facilitate genome wide analysis, a BAC library
containing 9216 clones with an average insert size of 130
kbp was constructed. This library represents greater
than 25X genome coverage. High density colony filters
are available upon request."

ORIGIN
  Query Match      7.0%; Score 559.8; DB 8; Length 620;
  Best Local Similarity 98.8%; Pred. No. 3.2e-119;
  Matches 564; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 609 GTTCCTTGGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTG 668
DB 1 GTTCCTTGGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTG 60
QY 669 GCGGAGTCCGGGCGAGGATCTCTGTCTATCTCACTTGTCTCCCTGCGCGAGAAAGTATCC 728
DB 61 GCGGAGTCCGGGCGAGGATCTCTGTCTATCTCACTTGTCTCCCTGCGCGAGAAAGTATCC 120
QY 729 ATCATGGCTGATGCAATCGCGGCTGCATAGCTTGATCCGGCTACCTGCCCATTCGAC 788
DB 121 ATCATGGCTGATGCAATCGCGGCTGCATAGCTTGATCCGGCTACCTGCCCATTCGAC 180
QY 789 CACCAAGCGAAACATCGCATCGAGCGAGCAGCTACTCGATGGAAGCGGCTGTGCGAT 848
DB 181 CACCAAGCGAAACATCGCATCGAGCGAGCAGCTACTCGATGGAAGCGGCTGTGCGAT 240
QY 849 CAGGATGATCTGACGAAGCATCAGGGGCTCGCGCCAGCGCACTGTTCGCCAGGCTC 908
DB 241 CAGGATGATCTGACGAAGCATCAGGGGCTCGCGCCAGCGCACTGTTCGCCAGGCTC 300
QY 909 AAGGCGCGCATGCCGACGCGAGGATCTCGTGTGACCCATGGCGATGCTGCTTGGCG 968
DB 301 AAGGCGCGCATGCCGACGCGAGGATCTCGTGTGACCCATGGCGATGCTGCTTGGCG 360
QY 969 AATATCATGGTGAATAATGGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTG 1028
DB 361 AATATCATGGTGAATAATGGCGCTTTCTGGAATTCATCGACTGTGGCGGCTGGGTGTG 420
QY 1029 GCGGACCGCTATCAGACATAGGTTGGCTACCGGTATATGCTGAAGGCTTGGCGGC 1088
DB 421 GCGGACCGCTATCAGACATAGGTTGGCTACCGGTATATGCTGAAGGCTTGGCGGC 480
QY 1089 GAATGGGCTGACCGCTTCTCTGCTGTTTACGGTATCGCGGCTCCCGATTCGACGCGCATC 1148
DB 481 GAATGGGCTGACCGCTTCTCTGCTGTTTACGGTATCGCGGCTCCCGATTCGACGCGCATC 540
QY 1149 GCCTTCTATCGCTTCTTGACGAGTCTTCT 1179
DB 541 GCCTTCTATCGGCTTCTTGACGAGTCTTCTT 571

RESULT 13
CK801630
LOCUS
DEFINITION
  CK801630 549 bp mRNA linear EST 26-FEB-2004
  NF32c08f44.r1 Tall Rescue PI283316 44 deg C Heat Stress SSH cDNA
  Schedonorus arundinaceus cDNA clone NF32c08f44 5', mRNA sequence.
```

```

CK801630
CK801630.1 GI:43400943
EST.
Schedonorus arundinaceus (Festuca arundinacea)
Schedonorus arundinaceus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Poae; Schedonorus.
1 (bases 1 to 549)
Zhang, Y., Zwonitzer, J. C., Chekhovskiy, K., May, G. D. and Mian, M. A. R.
A functional genomics approach for identification of heat tolerance
genes in tall rescue
(in) Hopkins, A., Wang, Z. Y., Mian, R., Sledge, M. and Barker, R. E.
(eds.), MOLECULAR BREEDING OF FORAGE AND TURF, Kluwer Acad. Pub. 0,
87-96 (2003)
Contact: Bruce A. Roe, University of Oklahoma, broe@ou.edu
Department of Chemistry and Biochemistry
Advanced Center for Genome Technology, University of Oklahoma
620 Parrington Oval, Norman, OK 73019, USA
Tel: 405 325 4912
Fax: 405 325 7762
Email: broe@ou.edu
Contact Dr. Rouf Mian (rmian@noble.org) regarding clone availability
Seq primer: M13 reverse primer
High quality sequence stop: 470.

FEATURES
  source      Location/Qualifiers
    1..549
      /organism="Schedonorus arundinaceus"
      /mol_type="mRNA"
      /cultivar="PI283316"
      /db_xref="taxon:4606"
      /clone="NF32c08f44"
      /tissue_type="shoot"
      /lab_host="E. coli"
      /clone_lib="Tall Rescue PI283316 44 deg C Heat Stress SSH
cDNA"
      /note="Vector: PCR 2.1-TOPO; Site 1: EcoR I; Site 2: EcoR
I; B/Clontech PCR-select cDNA subtraction library"

ORIGIN
  Query Match      6.8%; Score 542.4; DB 7; Length 549;
  Best Local Similarity 99.8%; Pred. No. 3.6e-115;
  Matches 543; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 556 TGAATGACTCGACGACGAGGAGGCGGCTATCGTGGCTGCGCAGCGGCGTTCCTT 615
DB 6 TGAATGACTCGACGACGAGGAGGCGGCTATCGTGGCTGCGCAGCGGCGTTCCTT 65
QY 616 GCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTGGCGAAG 675
DB 66 GCGCAGCTGTGCTCGACGTTGTCACTGAAGCGGGAAGGAGCTGGCTGCTATTGGCGAAG 125
QY 676 TCGCGGGGCGAGATCTCTGTGTCATCTACCTTGTCTCTCGCGAGAAAGTATCATCATGG 735
DB 126 TCGCGGGGCGAGATCTCTGTGTCATCTCCACCTTGTCTCTCGCGAGAAAGTATCATCATGG 185
QY 736 CTGATGCATGCGGGCTGTCATCGCTTGATCCGCTACCTGCGCATTCGACCCCAAG 795
DB 186 CTGATGCATGCGGGCTGTCATCGCTTGATCCGCTACCTGCGCATTCGACCCCAAG 245
QY 796 CGAAACATCGCATCGAGCGAGCAGCTACTCTCGATGGAAGCCGCTTGTGATCAGGATG 855
DB 246 CGAAACATCGCATCGAGCGAGCAGCTACTCTCGATGGAAGCCGCTTGTGATCAGGATG 905
QY 856 ATCTGGAAGAGCATCAGGGGCTCGCGCCAGCGAAGTGTTCGCGAGGCTCAAGGCGC 915
DB 306 ATCTGGAAGAGCATCAGGGGCTCGCGCCAGCGAAGTGTTCGCGAGGCTCAAGGCGC 365
QY 916 GCATGCCCGACGCGAGGATCTCGCTGACCCATCGCGATGCTCTCTGCGCAATATCA 975
DB 366 GCATGCCCGACGCGAGGATCTCGCTGACCCATCGCGATGCTCTCTGCGCAATATCA 425
QY 976 TGGTGGAAATGCCCGCTTTTCTGGATTATCATGACTGTGCGCGGCTGGGTGTGGCGGACC 1035
```

```

Db      426  TGGTGAATAAGCGCGCTTTCTCGATTTCATCGACTGTGGCCGGCTGGGTGTGGCGGACC 485
QY      1036  GCTATCAGACATAGCGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGCGAATGGG 1095
Db      486  GCTATCAGACATAGCGTTGGCTACCGGTGATTTGCTGAAGAGCTTGGCGCGAATGGG 545
QY      1096  CTGA 1099
Db      546  CTGA 549

RESULT 14
LOCUS   AQ447874
DEFINITION  mgxb0012101f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
ACCESSION  AQ447874
VERSION    AQ447874.1 GI:4577011
KEYWORDS   GSS.
SOURCE     Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM   Magnaporthe grisea
REFERENCE  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS    Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
           1 (bases 1 to 561)
           Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
           Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE      A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
           Genome
JOURNAL    Unpublished (1998)
COMMENT    Contact: Dean RA
           Clemson University Genomics Institute
           100 Jordan Hall, Clemson University, Clemson, SC 29634
           Tel: 864 656 5737
           Fax: 864 656 4293
           Email: rdean@clemson.edu
           Seq primer: TAATACGACTCATTATAGG
           Class: BAC ends
           High quality sequence start: 42
           High quality sequence stop: 326.
           Location/Qualifiers
             1..561
               /organism="Magnaporthe grisea"
               /mol_type="genomic DNA"
               /strain="70-15"
               /db_xref="taxon:148305"
               /clone="mgxb0012101f"
               /tissue_type="Protoplasts"
               /lab_host="E. coli DH10B"
               /clone_lib="CUGI Rice Blast BAC Library"
               /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
               Rice blast is one of the most devastating fungal diseases
               of rice world wide. It is a filamentous ascomycete with
               a haploid genome (n=7) of approximately 40 Mbp. Rice
               blast is an important model fungal pathogen for studying
               numerous aspects of the fungal-host interaction. In
               order to facilitate genome wide analysis, a BAC library
               containing 9216 clones with an average insert size of 130
               kbp was constructed. This library represents greater
               than 25X genome coverage. High density colony filters
               are available upon request."

ORIGIN
Query Match 6.4%; Score 514.8; DB 8; Length 561;
Best Local Similarity 99.6%; Pred. No. 9.9e-109;
Matches 516; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      510  CAGGGGCGCCGGTCTTTTGTCAAGACCGAGCTGTCCGGTGCCTGAATGAATCGAG 569
Db      44    CAGGGGCGCCGGTCTTTTGTCAAGACCGAGCTGTCCGGTGCCTTTTGAATCGAG 103
QY      570  GACGAGGCGAGCGGCTATCGTGGCTGGCCAGCGGCGGTCTCTTGGCGAGCTGTGCTC 629

```

```

Db      104  GACGAGGCGAGCGCGCTATCGTGGCTGGCCACGACGGGCGTTCCTTGGCGAGCTGTGCTC 163
QY      630  GACGTTGTCACTGAAGCGGGAAGGAGTGGTGTCTATTGGGCGAAAGTGGCGGGGAGGAT 689
Db      164  GAGGTTGTCACTGAAGCGGGAAGGAGTGGTGTCTATTGGGCGAAAGTGGCGGGGAGGAT 223
QY      690  CTCCTGTATCTACCTTGTCTCTCCGAGAAAGATATCCATCATGGCTGATGCAATGCGG 749
Db      224  CTCCTGTATCTACCTTGTCTCTCCGAGAAAGATATCCATCATGGCTGATGCAATGCGG 283
QY      750  CGGCTGCATACGCTTGTATCCGCTACCTGACCATTCGACCCCAAGCGGAATCATGCGATC 809
Db      284  CGGCTGCATACGCTTGTATCCGCTACCTGACCATTCGACCCCAAGCGGAATCATGCGATC 343
QY      810  GAGCGAGCAGCTACTTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTTGACCAAGAG 869
Db      344  GAGCGAGCAGCTACTTCGGATGGAAGCCGGTCTTGTGATCAGGATGATCTTGACCAAGAG 403
QY      870  CATCAGGGGCTCGGCCAGCGCAACTGTTCCGACAGGCTCAAGGCGCGCATGCCGACGGC 929
Db      404  CATCAGGGGCTCGGCCAGCGCAACTGTTCCGACAGGCTCAAGGCGCGCATGCCGACGGC 463
QY      930  GAGGATCTCGTCCGACCCATGGCGATCGCTGCTTGGCGAATATCATGTTGGAATGCG 989
Db      464  GAGGATCTCGTCCGACCCATGGCGATCGCTGCTTGGCGAATATCATGTTGGAATGCG 523
QY      990  CGCTTTCTCGGATTCATCGACTGTGGCGCGGTGGGTGT 1027
Db      524  CGCTTTCTCGGATTCATCGACTGTGGCGCGGTGGGTGT 561

RESULT 15
LOCUS   AQ447775
DEFINITION  mgxb0011E13f CUGI Rice Blast BAC Library Magnaporthe grisea genomic
ACCESSION  AQ447775
VERSION    AQ447775.1 GI:4576912
KEYWORDS   GSS.
SOURCE     Magnaporthe grisea (anamorph: Pyricularia grisea)
ORGANISM   Magnaporthe grisea
REFERENCE  Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
AUTHORS    Sordariomycetes incertae sedis; Magnaportheaceae; Magnaporthe.
           1 (bases 1 to 509)
           Yu, Y., Zhu, H., Boyd, C.A., Gaudette, B., Gayle, A., Kingsbury, R.,
           Phillips, K., Sasinowski, M., Wing, R.A. and Dean, R.A.
TITLE      A BAC End Sequencing Framework to Sequence the Magnaporthe grisea
           Genome
JOURNAL    Unpublished (1998)
COMMENT    Contact: Dean RA
           Clemson University Genomics Institute
           100 Jordan Hall, Clemson University, Clemson, SC 29634
           Tel: 864 656 5737
           Fax: 864 656 4293
           Email: rdean@clemson.edu
           Seq primer: TAATACGACTCATTATAGG
           Class: BAC ends
           High quality sequence stop: 342.
           Location/Qualifiers
             1..509
               /organism="Magnaporthe grisea"
               /mol_type="genomic DNA"
               /strain="70-15"
               /db_xref="taxon:148305"
               /clone="mgxb0011E13f"
               /tissue_type="Protoplasts"
               /lab_host="E. coli DH10B"
               /clone_lib="CUGI Rice Blast BAC Library"
               /note="Vector: pBACWICH; Site 1: HindIII; Site 2: HindIII;
               Rice blast is one of the most devastating fungal diseases
               of rice world wide. It is a filamentous ascomycete with

```

a haploid genome (n=7) of approximately 40 Mbp. Rice blast is an important model fungal pathogen for studying numerous aspects of the fungal-host interaction. In order to facilitate genome wide analysis, a BAC library containing 9216 clones with an average insert size of 130 kbp was constructed. This library represents greater than 25X genome coverage. High density colony filters are available upon request."

ORIGIN

```

Query Match      6.4%; Score 508; DB 8; Length 509;
Best Local Similarity 100.0%; Pred. No. 3.8e-107;
Matches 508; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 629 CGAGGTTGTACATGAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGC CGGGGACGGA 688
Db 1 CGAGGTTGTACATGAGCGGGAAGGACTGGCTGCTATTGGGCGAAGTGC CGGGGACGGA 60
Qy 689 TCTCCTGTCTATCACCCTTCTCTGCGGAGAAAGTATCCATCATGGCTGATGCAATGCG 748
Db 61 TCTCCTGTCTATCACCCTTCTCTGCGGAGAAAGTATCCATCATGGCTGATGCAATGCG 120
Qy 749 GCGGCTGCATACCGTTGATCCGGCTACTGCCCATTCGACCAACGCGAAACATCGCAT 808
Db 121 GCGGCTGCATACCGTTGATCCGGCTACTGCCCATTCGACCAACGCGAAACATCGCAT 180
Qy 809 CGAGCGAGCAGCTACTCGGATGGAAGCGGCTTGTGATCAGGATGATCTGGACGAAGA 868
Db 181 CGAGCGAGCAGCTACTCGGATGGAAGCGGCTTGTGATCAGGATGATCTGGACGAAGA 240
Qy 869 GCATCAGGGGCTCGCGCAGCCGAACTGTTCCGAGGCTCAAGGCGGCGCATGCCCGACGG 928
Db 241 GCATCAGGGGCTCGCGCAGCCGAACTGTTCCGAGGCTCAAGGCGGCGCATGCCCGACGG 300
Qy 929 CGAGGATCTCGTCCGTGACCCATGCGGATGCTGCTCCGAAATATCATGTGGAAAATGG 988
Db 301 CGAGGATCTCGTCCGTGACCCATGCGGATGCTGCTCCGAAATATCATGTGGAAAATGG 360
Qy 989 CCGCTTTTCTGGATTATCGACTGTGCGCGGCTGGGTGTGGCGGACCGCTATCAGGACAT 1048
Db 361 CCGCTTTTCTGGATTATCGACTGTGCGCGGCTGGGTGTGGCGGACCGCTATCAGGACAT 420
Qy 1049 AGCGTTGGCTACCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCT 1108
Db 421 AGCGTTGGCTACCGTGATATTGCTGAAGAGCTTGGCGGCGAATGGGCTGACCGCTTCCT 480
Qy 1109 CGTCTTTTACGGTATCCCGCTCCCGAT 1136
Db 481 CGTCTTTTACGGTATCCCGCTCCCGAT 508

```

Search completed: December 9, 2004, 03:36:21  
Job time : 21856 secs